Nature and Modulation of the Higher Order Chromatin Fibre

Nick Gilbert

Ph.D Thesis

Genes and Development Group
Department of Biomedical Sciences
The University of Edinburgh
2000

Declaration of Originality

I declare this thesis was written by me and is my own work.

Nick Gilbert

Acknowledgements

I would like to thank my two supervisors, Jim Allan and Austin Smith, for providing me with the opportunity to undertake my PhD in their labs. I am particularly grateful to Jim whose enthusiasm for the project was/is never-ending and who has spent many hours discussing my data and the finer points of chromatin structure with me.

I would like to thank everyone in the chromatin research group at George Square and Austin's lab at the CGR for being a good bunch to work and socialise with; especially members of Jim's lab: Colin, Chang-Hui, Melanie and Marietta.

Abstract

The bulk of eukaryotic cellular DNA is compacted into chromatin, a nucleoprotein complex that is responsible for packaging DNA into the nucleus and regulating gene transcription. The chromatin fibre is a dynamic structure which is able to accommodate the many complex processes which occur simultaneously in a living cell. The fundamental building blocks of the lower order chromatin fibre have been studied extensively, providing us with a detailed understanding of the structures present; much is known about how these structures are modulated to allow processes like gene transcription and replication to occur in an organised fashion. In contrast to our detailed knowledge of the fundamental building blocks of chromatin, little is known about the conformation of the higher order chromatin fibre. This lack of understanding is due, predominantly, to the inaccessibility of the higher order fibre for study, and that much of the research to date has considered the conformation of the higher order fibre to be uniform. In this project I have analysed and modulated the higher order chromatin fibre to assess the role this fibre plays in the regulation of cellular processes.

The conformation of the higher order chromatin fibre is often thought to change during the differentiation of cells. To study this alteration in conformation I have undertaken a detailed analysis of the higher order chromatin fibre from cells with different differentiation potentials and during their differentiation process. Using a hydrodynamic sedimentation approach to assess the conformation of the chromatin fibre I was unable to find any differences in its conformation. However, I have found that these chromatin fibres do have inherent differences in their nuclease sensitivities, suggesting that although the overall conformation of the fibres are similar, there are chromatin-related differences between cells with various differentiation potentials.

To study the uniformity of the higher order chromatin fibre at different chromosomal locations I have analysed the chromatin structure found at centromeres to determine whether there is an alteration in the conformation of the chromatin fibre, which might affect the function of centromeres. My results clearly show that in mouse and

human cells the chromatin fibre found at inner centromeric regions is more compact than the chromatin fibre present at outer centromeric regions, and in turn this is more compact than the bulk chromatin fibre. To determine the molecular basis for this, I have analysed acetylated centromeric heterochromatin from embryonic stem (ES) cells and heterochromatin associated with undermethylated centromeric DNA from F9 cells. My results demonstrate that this special chromatin architecture found at centromeres appears to be independent of histone acetylation and DNA methylation.

To establish whether an alteration in the chromatin conformation will alter a cell's differentiation potential I have expressed histone H5, a replacement linker histone normally found in nucleated erythrocytes, in pluripotential ES cells. My results show that the constitutive expression of H5 in ES cells causes substantial cell death. I have therefore constructed a regulated, tetracycline based, histone H5 expression system in ES cells, but I was unable to express H5 in a controlled manner to investigate the underlying chromatin structure of these cells. In addition, I expressed histone H5 DNA-binding mutants in ES cells which also caused substantial cell death. I was therefore unable to determine whether the cellular phenotype obtained from expressing H5 in ES cells was due to an alteration in chromatin structure or a non-specific effect from expressing a positively charged molecule. As a first step towards studying the expression of linker histones in living cells and during development, I constructed and analysed a green fluorescent protein (GFP)–histone H5 fusion. As for histone H5, the GFP-H5 fusion protein is correctly expressed in a variety of cell types, but is lethal to cells when expressed at high levels for longer periods of time.

Contents

Figures

Tables

List of Abbreviations

µg	1×10^{-6} grams
µl	1×10^{-6} litres
A_{260}	absorbance at 260nm
bp	base pair
BSA	bovine serum albumin
cfu	colony forming unit
DAPI	4,6-diamidino-2-phenylindole
ddNTP	dideoxynucleotide phosphate
DIA/LIF	leukemia inhibitory factor (Smith et al., 1988)
dNTP	deoxynucleotide triphosphate
DTT	dithiothreitol
EDTA	ethylene diamine-tetraacetic acid
EGTA	ethylene glycol-bis[β-aminoethyl ether]-tetraacetic acid
ES	embryonic stem
EtBr	ethidium bromide
g	acceleration or grams
hr	hour
IPTG	isopropyl β-D-thioglucoside
kb	kilo-base pair
kD	kilo-daltons
Mb	mega-base pair
MBA	3-methoxy benzamide
min	minutes
nt	nucleotide
°C	degrees centigrade
OD	optical density
PAGE	polyacrylamide gel electrophoresis
PBS	phosphate buffered saline
PMSF	phenyl methyl sulphonyl fluoride
SDS	sodium dodecyl sulphate
SSC	sodium chloride / sodium citrate buffer
TBE	tris / borate / EDTA buffer
TCA	trichloroacetic acid
TEMED	N,N,N',N'-tetramethyl-ethylenediamine
TEP	tris / EDTA / PMSF buffer
TPE	tris / phosphate / EDTA buffer
TST	tris / saline / tween buffer
V	volts
W	watts
X-Gal	5-bromo-4-chloro-3-indolyl β-D-galactopyranoside

1. Introduction

The genetic information which defines the characteristics of an organism is encoded by DNA. Cells package this DNA, using architectural and regulatory proteins, to confine and organise it into chromosomes, and to facilitate co-ordinated gene expression. Prokaryotes have evolved a variety of phylogenetically unrelated, small, basic, sequence-independent DNA-binding proteins (Sandman et al., 1998) that include histones in Euryarchaeota (Pereira and Reeve, 1998), and members of the HU family in many Bacteria (Oberto et al., 1994). In contrast, virtually, all Eukarya package their DNA using histone proteins (Van Holde, 1988). Higher eukaryotes have evolved larger, more complex genomes than prokaryotes which is critical for the complex patterns of gene expression required for the development of multicellular organisms. Consequently, this necessitates a higher level of DNA packaging and greater nuclear organisation.

There is great diversity in the size of eukaryotic genomes. A typical human cell contains approximately 6×10^9 base pairs of DNA which is unequally divided between 46 chromosomes (Figure 1.1) whilst the Puffer fish genome is only one-seventh the size of this (Mileham and Brown, 1994). Each double-helix in higher eukaryotes is packaged in an organised yet dynamic fashion such that the linear DNA is condensed some 10,000 times, giving a nuclear DNA concentration of 50-200mg/ml. In the early 1970's, DNA was found to be packaged into a nucleoprotein complex called chromatin, consisting of positively-charged histone and non-histone proteins which bind to the negatively-charged DNA and facilitate its folding into a higher order structure. In addition to packaging DNA, chromatin is responsible, in concert with many transcription factors, for regulating the expression of the 10,000 tissue-specific genes present in a typical vertebrate cell, of which only 10-20% are transcriptionally active at any particular time.

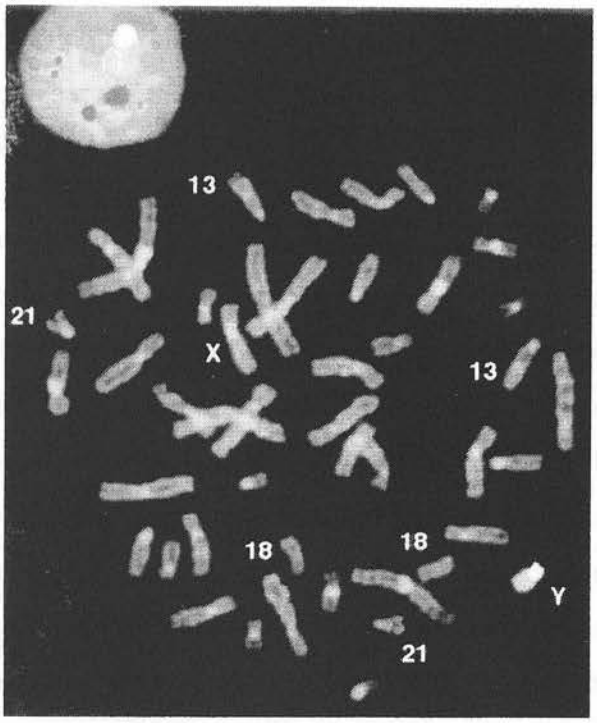


Figure 1.1 Human lymphocyte metaphase and interphase nuclei showing individual chromosomes. The nuclei are counter-stained with DAPI and are hybridised to centromeric probes for the X chromosome (lilac) and Y chromosome (yellow). Chromosome 13 is labelled using a YAC clone (green) and chromosome 21 is labelled by two cosmid clones (red) (Picture taken from Divane et al., 1994).

The more fundamental elements of eukaryotic chromatin architecture have been well studied. In contrast, the higher-levels of chromatin fibre organisation are poorly understood, although many researchers believe the fibre is wound into a solenoid structure and is arranged into a series of discrete loops which radiate from a scaffold-like structure. It is accepted that the chromatin fibre must have a highly dynamic architecture which permits cells to operate in a constantly changing environment, requiring the expression of many different genes. However, little is known about the different structures which the chromatin fibre can attain and, to date, it has not been shown whether there is a difference in the higher order chromatin fibre between one chromatin region and another.

Since the early days of chromatin biology the importance and robustness of the chromatin fibre has become increasingly apparent. In 1950, Stedman and Stedman wrote "It has always been a puzzle to us, how the physiological functions of the nuclei in the same organism can differ from one cell type to another when they all contain identical chromosomes and hence identical genes". This has now been explored using a variety of modern molecular techniques such that it is clear that growth, development and differentiation, in the main part, proceed through the regulated expression of genes part modulated by the conformation of the chromatin fibre. To further our understanding of the regulation of complex cellular processes it is necessary to examine in greater detail the interactions between the chromatin fibre and many other cellular components.

This introduction puts into context the studies to date which provide a background for my thesis and demonstrate both the complexity and importance of the chromatin architecture for organising cellular events and regulating gene expression. Finally, it provides a framework in which my studies of the higher order chromatin fibre can be related.

1.1 Chromatin structure

Eukaryotic cellular DNA is normally compacted into a 30nm chromatin fibre (Van Holde, 1988; Wolffe, 1995). The fundamental particle of this fibre is the nucleosome (Kornberg and Lorch, 1999) which consists of 160-200 base pairs of DNA wrapped approximately twice around an octamer of core histones and sealed with a single molecule of linker histone bound close to the core particle dyad (Thomas, 1999). Structural studies suggest that the 30nm fibre is composed of a tandem array of nucleosomes folded into a condensed higher order structure (Felsenfeld and McGhee, 1986; Van Holde and Zlatanova, 1996). The family of linker histones (including H1, H1° and H5) direct this folding (Thoma et al., 1979; Allan et al., 1981), and stabilise its conformation (Marion et al., 1985; Carruthers et al., 1998). Compacted chromatin fibres are generally repressive to DNA interactions and inhibit the movement of RNA polymerases (Kornberg and Lorch, 1991; Kornberg and Lorch, 1992) and the binding of transcription factors (Beato and Eisfeld, 1997) necessitating the remodelling of chromatin during transcription (Adams and Workman, 1993) and replication (Alexiadis et al., 1998).

The proteins associated with the chromatin fibre can be divided into two broad categories: histones and non-histones. The architectural histone proteins represent a well-studied group with unique compositions, sequence characteristics and functions which mediate the folding of DNA into chromatin and their expression is closely coupled to DNA synthesis (Wu and Bonner, 1981). Non-histone chromosomal proteins are essentially all of the other proteins isolated with chromatin, although in many cases these proteins act in concert with the histone proteins (Zlatanova and Van Holde, 1998). The family of histone proteins can be divided into core histones and linker histones. They both have similar amino-acid compositions and are often considered to be structural proteins, although in an increasing number of situations the role of linker histones is thought to be less structural and more regulating (Ohsumi et al., 1993; Shen et al., 1995) than the core histones.

1.1.1 Core histones

The main protein components of the chromatin fibre are the family of core histones. Core histones are positively charged proteins which bind directly to the DNA fibre by non-covalent forces, among which the electrostatic interactions between positively charged residues on the histones and DNA phosphates are probably the most important (Simpson et al., 1978). In higher eukaryotes there are four core histone proteins H2A, H2B, H3 and H4 which interact to form the core particle. Core histones are critical for the formation of the chromatin fibre and deletion of the H2B and H4 core-histones in yeast causes cells to arrest in G2 with concomitant loss of nucleosomes (Han et al., 1987; Kim et al., 1988).

All histones are basic proteins containing relatively large amounts of lysine and arginine (Stedman and Stedman, 1950) giving them a substantial net charge at physiological pH. This positive charge is unequally distributed and in all cases the N-terminal region contains a higher concentration of basic residues. All core histones have been remarkably conserved in length and amino acid sequence through evolution. Histones H3 and H4 are the most highly conserved with calf and pea histone H4 differing at only two sites in 102 residues (DeLange et al., 1969). Histones H3 and H4 have a central role both within the nucleosome (Arents et al., 1991) and in many chromosomal processes (for example, Juan et al., 1999) and these functional requirements presumably contribute to their sequence conservation.

Core histone proteins have three structural domains with the carboxyl-tail of the protein containing a 'histone-fold' motif (Arents et al., 1991). This motif is shared by many proteins, including a number of transcription factors (Xie et al., 1996; Hoffmann et al., 1996; Birck et al., 1998) and is thought to be involved in histone:histone and DNA:histone interactions. The 25-40 amino acid N-terminal domains of the core histones are highly conserved, indicative of their important role within the chromatin fibre. They have been shown by high resolution NMR to be mobile and flexible and they extend out of the nucleosome core in a manner that frees them for other interactions (Schroth et al., 1990). These amino terminal tails are

frequently modified and are thought to be critical in modulating the conformation of the chromatin fibre (Hansen et al., 1998).

1.1.2 Core histone variants

A number of core histone variants have been reported, many of which have occurred by gene duplication (Wang et al., 1996a; Wang et al., 1996b). The only core histone which does not appear to have any variants is H4. Histone variants are often expressed at specific stages during embryogenesis (Romano, 1992), or during cell differentiation (Bosch and Suau, 1995). Mouse histone variants have been well studied (for example Gunjan and Brown, 1999) although it is unclear to what extent these somatic variants modulate the conformation of the chromatin fibre. These variants fall into two main classes: those present in somatic cells, and a special set of 'testes' histones found only in developing spermatocytes. The testes-specific histone variants replace the somatic cell variants during the meiotic prophase of spermatogenesis. As the spermatocyte further develops into the mature sperm, all of these histones are replaced by protamines. The expression patterns of a number of the histones have been described: for example, in mouse testis H3.3A mRNA is present in pre- and postmeiotic cells, whereas expression of the H3.3B gene is essentially restricted to cells of the meiotic prophase (Bramlage et al., 1997). There are some more substantial sequence differences between histone variants: for example, the mouse H2B spermatid-specific histone has an extra 12 amino acids at its carboxyl terminus, but as it only constitutes 2% of total H2B protein it is unclear whether it has a functional role (Unni et al., 1995).

Recently, an alternative form of H2A, macroH2A, has been found associated with the inactive X chromosome of female mammals (Costanzi and Pehrson, 1998). The amino-terminal third of the mH2A protein is similar to full-length histone H2A, but the remaining two-thirds is unrelated to any known histones. This important observation links core-histone variants with a major alteration in the chromatin structure in a situation such as X-inactivation (Lyon, 1999). Recent data have shown that a conditional deletion of Xist (Brockdorff and Duthie, 1998) disrupts histone macroH2A localisation indicating that there might be a link between the Xist RNA,

which is nuclear matrix associated (Clemson et al., 1996), and macroH2A (Csankovszki et al., 1999). Another H2 variant, H2AX is phosphorylated on serine 139 when double-strand DNA breaks are introduced into mammalian cells by ionising radiation. H2AX foci then form at the sites of the DNA breaks suggesting a new mechanism for DNA damage detection and repair (Rogakou et al., 1999).

1.1.3 Core histone modifications

Post-translational modification of core-histone tails by acetylation, methylation, phosphorylation, ADP-ribosylation and ubiquitination (Wu et al., 1986; Davie and Chadee, 1998) are thought to be key arbiters of the chromatin fibre conformation and modulation (Wolffe and Hayes, 1999). These modification reactions tend to occur in the nucleus where proteins are also present to remove these covalent additions such that the turnover of the modifications can be very rapid. Acetylation is considered to be the modification most responsible for altering the conformation or stability of the chromatin fibre, probably by altering the secondary structure of the tails.

Two distinct types of core-histone acetylation occur *in vivo*. First, a number of histones are acetylated on serine residues in the cytoplasm. In most cells H1, H2A and H4 are found to contain such blocked N-termini (Pestana and Pitot, 1974). The second kind of acetylation is more important and occurs reversibly in the nucleus at specific groups of lysine residues on all the core histones but not the linker histones (Candido and Dixon, 1972). All of the acetylation sites are in the flexible N-terminal domains of the proteins suggesting that these sequences play some special role in chromatin organisation. The *in vivo* source of acetyl groups for histone acetylation seems to be acetyl coenzyme A, and the transfer of acetyl groups to histone lysine residues is catalysed by a number of acetyltransferases. The acetylation of histones in nuclei is a dynamic process and enzymes exist that catalyse the deacetylation reaction (Ayer, 1999). Sodium butyrate (Kruh, 1982) has been shown to be an inhibitor of deacetylases and addition of low levels of butyrate to cultured cells gives rise to highly acetylated forms of the histones (Riggs et al., 1977) making it an important tool for studying histone acetylation. In addition, histone deacetylases can be inhibited by two specific fungal toxins (trichostatin A and trapoxin) which are now

used more frequently than sodium butyrate as they do not appear to have any side-effects on other cellular processes (Yoshida et al., 1995). Amphibian embryos grown in high concentrations of trichostatin A, sufficient to cause constitutive hyperacetylation, will develop up to gastrulation whereas later development is severely disrupted, indicating the importance of histone acetylation for the regulation of gene expression (Almouzni et al., 1994).

H3 is pivotal in the organisation of the chromatin fibre. H3 is partially replaced at centromeres by two related proteins including CENP-A and CSE-4 (Section 1.3.2) which may alter the conformation of the fibre. H3 is phosphorylated at serine 10 and serine 28 in a cell cycle dependent manner (Shibata et al., 1990; Wei et al., 1999; Goto et al., 1999), possibly altering its binding in mitotic cells through a re-arrangement of its amino-terminal domain (Mazen et al., 1987; Sauve et al., 1999). Finally, H3 is methylated by CARM1, a protein methyltransferase, possibly as a mechanism to regulate transcriptional regulation, which then enhances transcriptional activation of nuclear receptors (Chen et al., 1999).

1.1.4 Structure and function of nucleosomes

The nucleosome is the fundamental repeating unit of chromatin (Kornberg and Lorch, 1999). It was discovered through its appearance in the electron microscope as the beads in the extended beads-on-a-string structure of decondensed chromatin fibres (Thoma and Koller, 1977) (Figure 1.2), by the generation of approximately 200 base pair DNA ladders upon digestion of chromatin with endogenous or exogenous nucleases (Noll, 1974) and by the isolation of 11.5S nucleoprotein complexes (Sahasrabuddhe and Van Holde, 1974). Histone/histone cross-linking patterns and X-ray diffraction data led to a working model of the nucleosome (Kornberg, 1974; Kornberg and Thomas, 1974), which was later confirmed by a low-resolution crystal structure of the core particle (Klug et al., 1980), and then solved to 7Å resolution (Richmond et al., 1984). This data suggested that each nucleosome was a complex consisting of DNA and histones, in a stable conformation at physiological ionic strength (Germond et al., 1976).

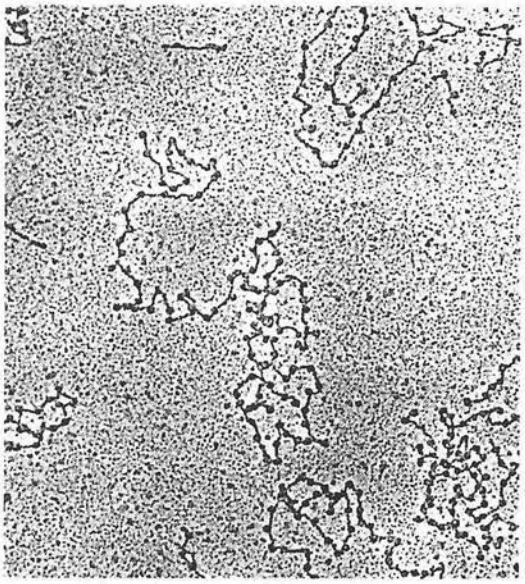


Figure 1.2 Nucleosomes are visible in decondensed chromatin fibres. Chicken erythrocyte chromatin was depleted of linker histones, resuspended at low ionic strength and visualised by electron microscopy (Picture supplied by J. Allan).

Within the nucleosome core, 146 base pairs of DNA are wrapped 1.75 times around the histone octamer in a left-handed superhelix with about 7.6 turns of the DNA helix per superhelical turn. Each octamer is tripartite and consists of a central tetramer of histones H3/H4 flanked by two H2A/H2B dimers (Arents et al., 1991). The nucleosome proper consists of a histone octamer, a linker histone and the linker DNA between nucleosomes with over 80% of cellular DNA incorporated into nucleosomes (Noll, 1974). In addition, a subnucleosomal particle has been defined as the chromatosome which consists of the nucleosome core, a linker histone and 168 base pairs of DNA (Simpson, 1978a). A high resolution structure of the nucleosome core particle has recently been determined at 2.8Å by crystallising the core particle in the presence of a defined 146 base pair palindromic DNA sequence (Luger et al., 1997). This shows in detail the interacting bonds between the core particle and specific DNA sequences (Widom, 1997), although the unstructured histone tails are still not clearly visible.

The nucleosome is capable of forming two distinct functions. Firstly, as a structural component of the chromatin fibre and secondly, as a modulator of gene expression. Nucleosomes are often considered to be inhibitory to gene expression and because of this it is necessary for their positions on the DNA to be modulated to allow gene expression to occur. Nucleosomes have been shown to be precisely positioned (Simpson, 1991) across genes (Davey et al., 1995) in a manner which allows key transcription factor binding sites to be exposed but for the fibre to be folded in a stable conformation. The positioning of nucleosomes can be modulated by minor alterations in the DNA sequence and by DNA methylation (Davey et al., 1997) presumably due to specific interactions between the DNA sequence and the surface charges on the nucleosome. Nucleosome positioning has been shown to differ between the *Xenopus* oocyte and somatic 5S RNA sequences, modulating the efficiency of H1 binding and thus TFIIIA transcription (Panetta et al., 1998). Although histones are positioned discretely they are mobile (Meersseman et al., 1992) and this is thought to be important for the remodelling of genes, to allow transcription to proceed.

The development of defined nucleosome arrays has provided a model system for exploring the conformation of the chromatin fibre in a controlled environment (Fletcher and Hansen, 1996). Nucleosome arrays lacking the core histone N-termini can neither fold (Garcia-Ramirez et al., 1992; Fletcher and Hansen, 1995) nor oligomerise (Carruthers et al., 1998), indicating the importance of the tails for higher order structure formation. High salt can be used to substitute for the tail domains and induce folding, by inducing the tails to move to a non-nucleosomal location (Fletcher and Hansen, 1995). Core histone acetylation significantly enhances transcription factor access and transcription on nucleosomal arrays (Lee et al., 1993; Vettese-Dadey et al., 1996; Tse et al., 1998), but this effect is inhibited by the addition of linker histones (Ura et al., 1997). This suggests that acetylation of the core histones causes an as yet undefined alteration in the conformation or stability of the nucleosome. Nucleosome mobilisation can be catalysed by the SWI/SNF complex (Whitehouse et al., 1999) which may be mediated via the H2B N-terminus *in vivo* (Recht and Osley, 1999) although SWI/SNF has been shown to function in the absence of the histone tails (Guyon et al., 1999). The interaction between the NURF complex and the nucleosome is also inhibited if the N-terminal histone tails are removed, further suggesting a functional link between these complexes and the N-terminal core histone tails (Georgel et al., 1997),

1.1.5 Nucleosome assembly

During replication parental nucleosomes are segregated to the daughter DNA strands in a dispersive manner (Krude and Knippers, 1991) requiring continuous nucleosome production and assembly (Adams and Kamakaka, 1999). After replication H3/H4 tetramers bind to the DNA followed by H2A/H2B dimers. Newly synthesised H3 and H4 are acetylated at lysine 5 and 12 in *Drosophila* and HeLa cells (Sobel et al., 1995), and when replication occurs in the presence of sodium butyrate the chromatin fails to form (Perry and Annunziato, 1989); although histone deacetylation is required for the maturation of newly replicated chromatin (Annunziato and Seale, 1983).

Newly synthesised H3 and H4 are assembled onto replicating DNA *in vitro* by Chromatin assembly factor 1 (Verreault et al., 1996; Chang et al., 1997). CAF-1 recognises the H3/H4 histone complex if at least one H4 site is acetylated, serving as a chemical tag or changing the protein conformation to allow CAF-1 to assemble the H3/H4 complex. There is some redundancy in the H3 and H4 amino termini for the assembly of the H3/H4 tetramer; either can be deleted without preventing nucleosome assembly in yeast cell extracts or in *Xenopus* cell extracts, but deletion of both in the same extract blocks assembly (Ling et al., 1996). CAF-1 is also responsible for the eipigenetic maintenance of chromosome states *in vivo* by interacting with PCNA, a DNA polymerase clamp, which marks replicated DNA (Shibahara and Stillman, 1999).

The mechanism whereby H2A and H2B are assembled is less clear. NAP-1 has been identified in humans, *Drosophila* and yeast and facilitates the assembly of nucleosomal arrays in an ATP-dependent manner (Ito et al., 1996). This factor is bound to H2A and H2B in whole-cell extracts and moves from the nucleus in S phase (when most of the assembly is taking place) to the cytoplasm in G2. *In vivo* cell labelling studies indicate that there is a two-stage maturation process for newly replicated chromatin (Smith et al., 1984). The H3/H4 tetramers are, initially, weakly associated with newly replicated DNA and only later become stably associated with the fibre. In contrast more than 50% of newly synthesised H2A and H2B, and essentially all of the new H1, are deposited at sites on the bulk chromatin distinct from material containing newly replicated DNA (Jackson et al., 1981).

1.1.6 Linker histones

Linker histones (H1, H5, H1°) are located on the outside of the nucleosome (Pruss et al., 1995) and are responsible for the condensation of the chromatin fibre (Thoma and Koller, 1977; Thoma et al., 1979). The molecular interactions which associate linker histones with the chromatin fibre and induce higher order folding are poorly understood. Replacement linker histones (H1° and H5) have been implicated in stabilising the chromatin fibre (Marion et al., 1985; Sun et al., 1990c) more than H1. Linker histones are constructed from a highly conserved 79-residue central globular domain with flexible, flanking N and C-terminal domains (Hartman et al., 1977; Aviles et al., 1978). The NMR (Clore et al., 1987) and crystal structures (Ramakrishnan et al., 1993) of the globular domain of H5 have been obtained and show a high degree of similarity with the DNA binding protein CAP (McKay and Steitz, 1981; Schultz et al., 1991), RAP30 (Groft et al., 1998) and the transcription factor HNF-3 (Clark et al., 1993).

The globular domain of the linker histone directs its localisation on the nucleosome (Allan et al., 1980; Allan et al., 1981) in a process which is modulated by both its tail domains and by interactions with the core histones (Allan et al., 1986). At physiological pH the globular domain remains bound to the nucleosome leaving the C- and N-terminal domains free to interact with linker DNA. This probably influences the entry/exit angle of linker DNA which may in turn dictate the extent to which the higher order fibre can be folded (Bednar et al., 1998). Due to the interaction of the linker histone with the linker DNA and its role in the folding of condensed and transcriptionally inactive chromatin it is necessary to identify the exact position of the globular domain and extended tail domains of the linker histone on the nucleosome to understand how the higher order chromatin fibre is folded (Crane-Robinson, 1997; Travers, 1999).

Early models indicated that the globular domain of the linker histone bound to the central gyre at the nucleosomal pseudodyad (Figure 1.3Aii) (Allan et al., 1980) whilst neutron diffraction of chromatosomes (Lambert et al., 1991) support an asymmetric position for the globular domain of H5 (Figure 1.3Aiii), spanning the

gyres between one terminus and the midpoint. An alternative asymmetric model, derived from studies of a chromatosome reconstituted on *Xenopus* 5S ribosomal DNA, proposed that the globular domain of H5 makes a single contact with DNA on the inside of the DNA gyre (Figure 1.3Aiv) (Pruss et al., 1996; Hayes, 1996). Recent cross-linking studies have conclusively mapped the binding site of the globular domain of linker histone H5 and showed that in bulk chromatin the globular domain of the linker histone forms a bridge between one terminus of chromatosomal DNA and the midpoint of the nucleosome (Zhou et al., 1998) (e.g. Figure 1.3Aii, 1.3Aiii).

Another approach which has been used for identifying the position of the linker histone on the nucleosome is to determine the protection it can confer to the DNA from nuclease digestion (Figure 1.3B). This approach may also indicate the position of the linker histone tails although the globular domain of H1 can almost give the same protection of nucleosomal DNA against micrococcal nuclease as can full-length H1 (Allan et al., 1980). An early study used DNaseI foot-printing to demonstrate that the globular domain of H1/H5 is responsible for protection at the dyad axis, whilst the extended N- and C-domains are thought to be responsible for partial protection at sites further away from the dyad axis (Staynov and Crane-Robinson, 1988). Other studies have suggested that the linker histone protects the chromatosome asymmetrically although it is now thought that the sequence used for these experiments may be quite unique in its ability to position nucleosomes (Hayes and Wolffe, 1993). A recent comprehensive study from our laboratory (Shen, 1997) analysed a number of nucleosomes and found that in many situations the chromatosome was symmetrically protected, however for some nucleosomes the protected fragment was asymmetrically located to various extents. This suggests that although there is a tendency for the linker histone to sit symmetrically at the dyad axis, at specific sequences the linker histone can sit asymmetrically.

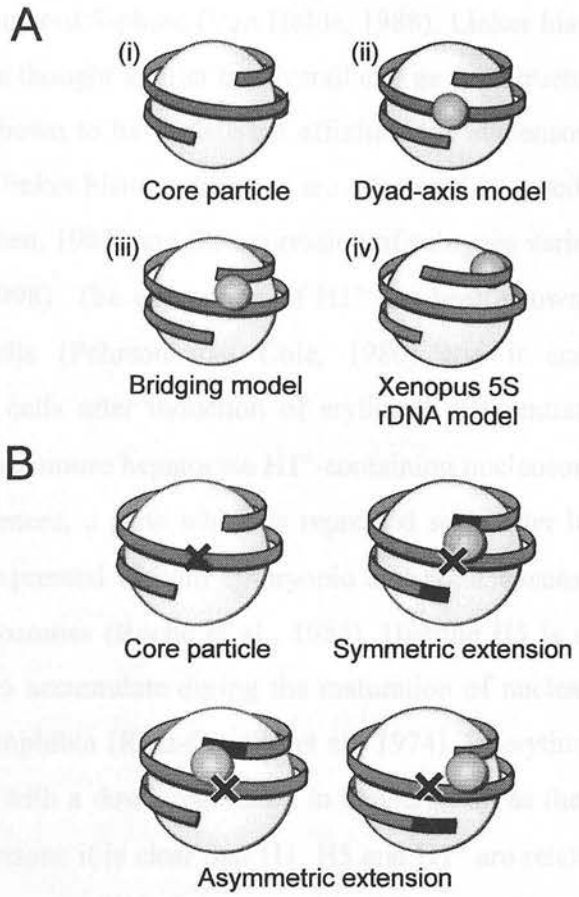


Figure 1.3 Position of linker histone globular domain on the nucleosome core particle. (A) Models for the placement of a linker histone globular domain on the nucleosome core particle. (B) Possible locations of the linker histone globular domain on the nucleosome to give different extents of nuclease protection. X, histone-octamer dyad. (Picture taken from Travers, 1999).

1.1.7 Linker histone variants

There are six mouse histone H1 genes present in the histone gene cluster on mouse chromosome 13, that encode the somatic H1s denoted histones H1a to H1e, and the testis specific H1t. The H1° gene is located on mouse chromosome 15 (Wang et al., 1997). Linker histones are physically less stable than core histones and are synthesised throughout S-phase (Van Holde, 1988). Linker histone variants and their modifications are thought to alter the overall charge and structure of the nucleosome, and have been shown to have different affinities for nucleosomal arrays (Ura et al., 1996). Different linker histone subtypes are expressed in specific cell types in mouse (Lennox and Cohen, 1983) and the expression of subtypes varies during development (Clarke et al., 1998). The expression of H1° has been shown to be upregulated in differentiated cells (Pehrson and Cole, 1980) and it accumulates in murine erythroleukemia cells after induction of erythroid differentiation (Rousseau et al., 1991). In addition, mouse hepatocyte H1°-containing nucleosomes are enriched in α-fetoprotein sequences, a gene which is repressed soon after birth, and the albumin gene which is expressed in both embryonic and adult tissues is absent from H1°-containing nucleosomes (Roche et al., 1985). Histone H5 is a replacement histone which is found to accumulate during the maturation of nucleated erythrocytes from birds, fish and amphibia (Ruiz-Carrillo et al., 1974). In erythrocytes, the expression of H5 correlates with a down regulation in transcription as the cells enter G0. From sequence comparisons it is clear that H1, H5 and H1° are related although a number of observations suggest that they may play different roles in the architecture and regulation of the chromatin fibre (Crane-Robinson, 1999). *Saccharomyces cerevisiae* has an unusual non-essential linker histone, containing two globular domains, deletion of which has no detectable effect on gene expression (Landsman, 1996; Ushinsky et al., 1997; Patterton et al., 1998). *Tetrahymena* histone H1, unlike metazoan H1s, lacks a globular domain. Deletion of the gene does not influence expression from the majority of genes, however a subset of genes are either activated or repressed in H1-deficient strains (Shen and Gorovsky, 1996). Also a single copy of a linker histone gene is enough for proliferation of the DT40 chicken B cell line (Takami and Nakayama, 1997), and mouse H1° is dispensable (Sirotkin et al., 1995),

suggesting that linker histones have redundant functions. The prevailing idea is that histone H1 can be considered as a gene-specific regulatory protein rather than an architectural protein (Wolffe, 1999; Crane-Robinson, 1999), although other classes of linker histones such as H1° probably have a more general repressive role.

1.1.8 Linker histone modifications

Linker histones contain sites of amino acid phosphorylation, many of which have cdc2/cdk2 consensus sequences. They are extensively phosphorylated in proliferating cells in a cell-cycle dependent manner such that phosphorylation is highest in M phase when chromatin is most condensed (Bradbury, 1992). During mitosis, genomic DNA is packaged into condensed chromosomes to facilitate its accurate segregation to daughter cells. Concomitant with mitotic chromosome condensation histone H1 is highly phosphorylated (Hohmann, 1983) which is thought to help compact the 10nm fibre into a 30nm fibre (Dolby et al., 1979; D'Anna et al., 1979). Phosphorylation has been shown directly to weaken the interactions of the basic tails of the linker histone with DNA possibly destabilising the chromatin fibre (Hill et al., 1991). Reconstitution experiments onto SV40 minichromosomes with differentially phosphorylated forms of histone H1 clearly indicate that H1 phosphorylation affects the chromosome structure and the minichromosomes replicate with different efficiencies (Halmer and Gruss, 1996). H1 phosphorylation has also been linked to alterations in the chromatin fibre such as GR-mediated disruption of the MMTV promoter *in vivo* (Lee and Archer, 1998) and H5 phosphorylation is known to inhibit the condensation of the chromatin fibre in chicken erythrocytes (Sung, 1977) and when ectopically expressed in other cell types (Aubert et al., 1991). However, some data suggests that histone phosphorylation is unnecessary for chromatin condensation. For example, in a *Xenopus* mitotic extract depleted of H1 the structural reorganisation that leads to condensed metaphase chromosomes is still active (Ohsumi et al., 1993) and H1 phosphorylation in *Tetrahymena* is not essential for cell viability (Mizzen et al., 1999). Also, in *Tetrahymena*, macronuclei H1 is phosphorylated throughout the cell cycle although phosphorylated and non-

phosphorylated forms of H1 are localised to different chromatin domains suggesting they may be involved in regulating gene transcription (Lu et al., 1995).

1.1.9 HMG proteins

HMG (high mobility group) chromosomal proteins are one of the most abundant class of non-histone chromosomal protein. These proteins have been grouped together based on their ubiquitous distribution in the nuclei of all higher eukaryotes, on their association with isolated chromatin, on common physical properties and on the observation that all of these proteins act as architectural elements that affect DNA-related activities in the context of chromatin. The HMG chromosomal proteins can be split into three groups: HMG 1 and 2, HMG 14 and 17 (Bustin et al., 1995) and the HMG I proteins which include HMG-I, HMG-Y and HMG-C. The three classes of HMG protein each have their own functional motif (Bustin, 1999). HMG-1/-2 proteins have an HMG domain consisting of an L-shaped arrangement of three α-helices, with two independent DNA-binding surfaces on the 'L' at right angles to each other. The HMG-14/-17 proteins have a nucleosomal binding motif and the HMG-I/Y group has an AT hook motif. Although the HMG proteins are relatively abundant the amount of HMG-1/-2 in a cell is about 10-fold lower than that of a histone, the amount of HMG-14/-17 is 10-fold less than HMG-1/-2 and HMG-I/Y is 10-fold lower than that of HMG-14/-17. Therefore, the amount of HMG-14/-17 in an average cell is only enough to bind to 1% of nucleosomes so even small fluctuations in the levels could have a major effect on transcription.

HMG-1 binds linear DNA with moderate affinity and no sequence specificity, but does bend the double helix significantly on binding through the minor groove. HMG-1 is implicated in regulating the association of transcription factors with DNA. HMG-2 is expressed at a much lower level that HMG-1 (Bustin and Reeves, 1996) whilst HMG-4 is almost exclusively expressed in embryos (Vaccari et al., 1998). HMG1 mouse knockouts survive until shortly after birth when they die from lethal hypoglycaemia (Calogero et al., 1999). HMG-1 is not essential for the overall organisation of chromatin in the cell nucleus but is critical for proper transcriptional control by specific transcription factors. Cell lines lacking HMG-1 grow normally,

but the activation of gene expression by the glucocorticoid receptor, and presumably other genes, is impaired (Calogero et al., 1999).

HMG-14 and HMG-17 bind specifically to nucleosome core particles and are thought to alter the conformation of the chromatin fibre. HMG-14 alleviates histone H1-mediated inhibition of transcription by RNA polII and enhances transcription on chromatin templates (Ding et al., 1997). Confocal microscopy has been used to localise HMG-14/-17 to specific chromosomal domains (Postnikov et al., 1997) and HMG-17 has been localised to sites of active transcription; in the absence of transcription it is relocated to inter-chromatin granules (Hock et al., 1998). HMG-17 is acetylated by PCAF which alters its interaction with nucleosomes (Herrera et al., 1999a), and HMG-14 is phosphorylated on serine 6 by different kinases, depending on the stimulus, providing a functional link between cell surface events and the chromatin fibre (Thomson et al., 1999). The N-terminal domain of HMG-14 interacts with histone H2B, whilst the C-terminal region, which is responsible for the transcriptional enhancement and chromatin unfolding activities of HMG-14, interacts with the N-terminus of histone H3 (Trieschmann et al., 1998). The N-terminal regions of the core histones are important for the folding of oligonucleosomes (Allan et al., 1982a; Garcia-Ramirez et al., 1992; Fletcher and Hansen, 1995) and the interaction of HMG-14 with the H3 tail may loosen the conformation of the chromatin fibre. Gene deletion studies in chicken DT-40 cells have shown that neither HMG-14 nor HMG-17 are required for cell viability (Li et al., 1997). Although, as found for the deletion of HMG-1 (Calogero et al., 1999), the deletion of HMG-14/-17 may only effect the expression of a specific class of genes such that a phenotype may only become apparent during a specific stage in development.

All the HMGI proteins bind specifically to A-T rich runs of double-stranded DNA (Solomon et al., 1986) and *in vitro* bind to the narrow minor groove by means of an 11-amino acid binding domain known as the 'A-T hook motif' (Reeves and Nissen, 1990). The three HMGI proteins are present in rapidly dividing cells and are virtually absent in terminally differentiated non-dividing cells (Johnson et al., 1988). These proteins have been correlated with a neoplastic (Giancotti et al., 1993) and /or

undifferentiated phenotype (Johnson et al., 1988; Johnson et al., 1989; Johnson et al., 1990) suggesting they may be involved in maintaining the undifferentiated state of chromatin. The role of HMGI proteins is a little confusing as they also bind to multiple sites in mouse satellite DNA suggesting they may be involved in the condensation of centromeric heterochromatin (Radic et al., 1992).

1.1.10 30nm chromatin fibre

Most eukaryotic interphase chromosomes are packaged in the form of a folded 30nm fibre (Langmore and Schutt, 1980). The conformation of the fibre was originally characterised by electron microscopy, using either chromatin which had been unfolded under low salt conditions (Oudet et al., 1975) or by the removal of the linker histones (Thoma and Koller, 1977) to give a 'beads-on-a-string like-structure'. The condensation of the chromatin fibre is primarily dependent on electrostatic interactions (Clark and Kimura, 1990) as shown by the folding of the extended 10nm fibre to the compact 30nm fibre by the addition of mono- or divalent cations (Thoma et al., 1979; Carruthers et al., 1998). *In vivo*, the fibre is compacted by the addition of linker histones which are able to reduce the electrostatic free energy of the fibre by displacing bound cations and reducing the residual charge. The role of the specific domains of H1 responsible for this compaction are unclear especially as it is uncertain where the linker histone tail domains bind in the nucleosome. Early studies demonstrated that the carboxyl terminus of H1 was required for chromatin folding (Allan et al., 1986) whilst more recent microscopic studies suggest the tails are not as important as the globular domain in forming a stable fibre (Leuba et al., 1998a; Leuba et al., 1998b).

Each cellular system has a characteristic average repeat length of linker DNA which reflects the nucleosome repeat length of chromatin i.e. the distance between nucleosomes (Van Holde, 1988). The nucleosome repeat length has been correlated to gene expression with the repeat length increasing during erythropoiesis in chicken (Weintraub, 1978), or a reduction in the repeat length during sea urchin embryogenesis and a short repeat length in calf cortical neurons during differentiation (Allan et al., 1984). The ability to produce different repeat lengths

depends not only on the balance between the amounts of core histone octamers and linker histones but also on the total histone/DNA ratio (Stein and Mitchell, 1988). For example, the nucleosome repeat length of tissue culture cells is shorter than that of primary cells as cells growing in culture are consistently selected for growth and transcription of the histone genes is probably maximal. In addition, repeat lengths are determined by electrostatic interactions in the nucleosome such that the folding of the fibre is dependent on the neutralisation of charges in the linker DNA (Blank and Becker, 1995). However, the replacement of histone H1 by H5 *in vivo*, which would be expected to alter the electrostatic interactions, does not change the nucleosome repeat length but does increase the fibres stability (Sun et al., 1990c). In addition, when exogenous H1 is expressed in *S. cerevisiae* it binds to chromatin but does not alter the nucleosome spacing (Linder and Thoma, 1994).

The conformation of the higher order fibre is determined by a number of factors in addition to the dispersion of charges by the linker histones. The extended tail domains of both the linker histones and the core histones interact as shown by cross-linking studies (Boulikas et al., 1980). Removal of the core histone tails abolishes the normal folding of the chromatin fibre in the presence of linker histones (Allan et al., 1982a) or during salt-dependent compaction (Garcia-Ramirez et al., 1992). This suggests that the core histone tails are responsible for locating the linker histone tails in a position where they are able to shield the electrostatic interactions required for fibre condensation. It also demonstrates how core histone acetylation may play a role in the dissociation of the linker histone tails in a manner which could destabilise the fibre. However, studies have been unable to demonstrate that hyperacetylated chromatin has a different conformation to hypoacetylated chromatin (McGhee et al., 1983a).

The conformation of the 30nm fibre is controversial (Van Holde and Zlatanova, 1996; Ramakrishnan, 1997) and many different models for its structure have been proposed (Figure 1.4). There are three classes of helical model: the solenoid, the twisted ribbon and the crossed-linker models and one non-helical 'superbead' model (Hozier et al., 1977; Renz et al., 1977; Zentgraf and Franke, 1984). In the solenoid

model (Figure 1.4A), the nucleosomes are organised into a helical array with about 6 to 8 nucleosomes per turn with histone H1 on the inside of the filament (Finch and Klug, 1976; Thoma et al., 1979; Graziano et al., 1994). In this model, the linker DNA is assumed to be on the inside of the solenoid and has to be bent to connect each nucleosome with its neighbour along the helical filament. In a variant of this model (Figure 1.4B), on the basis of electric dichroism studies, it is proposed that the linker DNA is bent in a continuation of the nucleosomal DNA (McGhee et al., 1983b). Both of these models are notably characterised by having bent linker DNA whereas the helical ribbon model (Figure 1.4D) (Woodcock et al., 1984), suggests that nucleosomes form a zigzag with the linker DNA being largely straight and parallel to the fibre axis with H1 maintaining the entry/exit angle of the linker DNA into the nucleosome (Yang et al., 1994). In the crossed-linker model (Figure 1.4D) DNA is thought to criss-cross between nucleosomes in a direction roughly perpendicular to the filament axis with H1 located on the inside of the fibre (Williams et al., 1986). The effect of heterogeneity in linker length and the relative rotation of the nucleosomes has been analysed by model building studies and compared to microscopic images resulting in an irregular model for the chromatin fibre (Figure 1.4E) (Woodcock et al., 1993; Horowitz et al., 1994; Woodcock, 1994; Leuba et al., 1994). Recent experimental data (Bednar et al., 1998), suggest that there is a linker histone-dependent architectural motif beyond the level of the nucleosome core particle that takes the form of a stem-like structure generated by the entry and exit linker DNA segments into the nucleosome. The stem motif directs the arrangement of nucleosomes and linker DNA within the chromatin fibre establishing a unique three-dimensional zigzag folding pattern which has also been identified using chromatin fragmentation in living cells (Rydberg et al., 1998).

The key difference between these models for the conformation of higher order chromatin fibre is a topological one. If it were possible to determine the linking number of the higher order fibre it would, in all probability, demonstrate the formation of a specific fibre. However, in the absence of knowing the linking number it is only by examining indirect evidence which can give us an insight into the conformation of the higher order fibre. In the solenoidal, or sequential, class of

models each nucleosome is connected to its neighbour along the helical array by linker DNA. However, in the nonsequential models such as the helical ribbon and crossed-linker models, nucleosomes are connected to each other diametrically across the filament or to the nucleosome below rather than to their immediate neighbours, thus having a different linking number. This may be significant as the fibre's linking number may affect its propensity to modulate its dynamic conformation. Two main areas in which the sequential and non-sequential models differ is the mass per unit length of the compacted fibre, and the structure of the linker DNA i.e. whether it is straight or bent. On both of these points the experimental data are contradictory thus not providing an unambiguous conformation for the fibre.

No single model of the 30nm fibre fits all of the evidence, but two important pieces of data strongly support the solenoid model. The first is that chromatin with either very short linkers (Allan et al., 1984; Pearson et al., 1984) or very long linkers, as in sea urchin sperm (Widom et al., 1985), form a 30nm fibre rather than the fibre diameter being dependent on linker length. Secondly, increased stabilisation is seen when over six nucleosomes form part of a polynucleosome chain (Butler and Thomas, 1980) or when six contiguous nucleosomes in long chains have H5 (Graziano et al., 1988) or H1 bound (Goytisolo et al., 1996). In the solenoid model, six nucleosomes form a full turn, which can then be stabilised because the two nucleosomes at the ends are one pitch of the helix apart and come into contact. In other models it is difficult to explain why any discontinuity in behaviour should occur at six nucleosomes. A number of studies by Yao et al. (1990a; 1991) have also shown that the linker DNA is bent although recent cryomicroscopy data does indicate the linker DNA is straight favouring the irregular structure (Figure 1.4E). Finally, the recent high resolution crystal structure of the nucleosome core particle (Luger et al., 1997), in conjunction with model building studies, does suggest that the structure of the nucleosome core particle is compatible with the solenoid model.

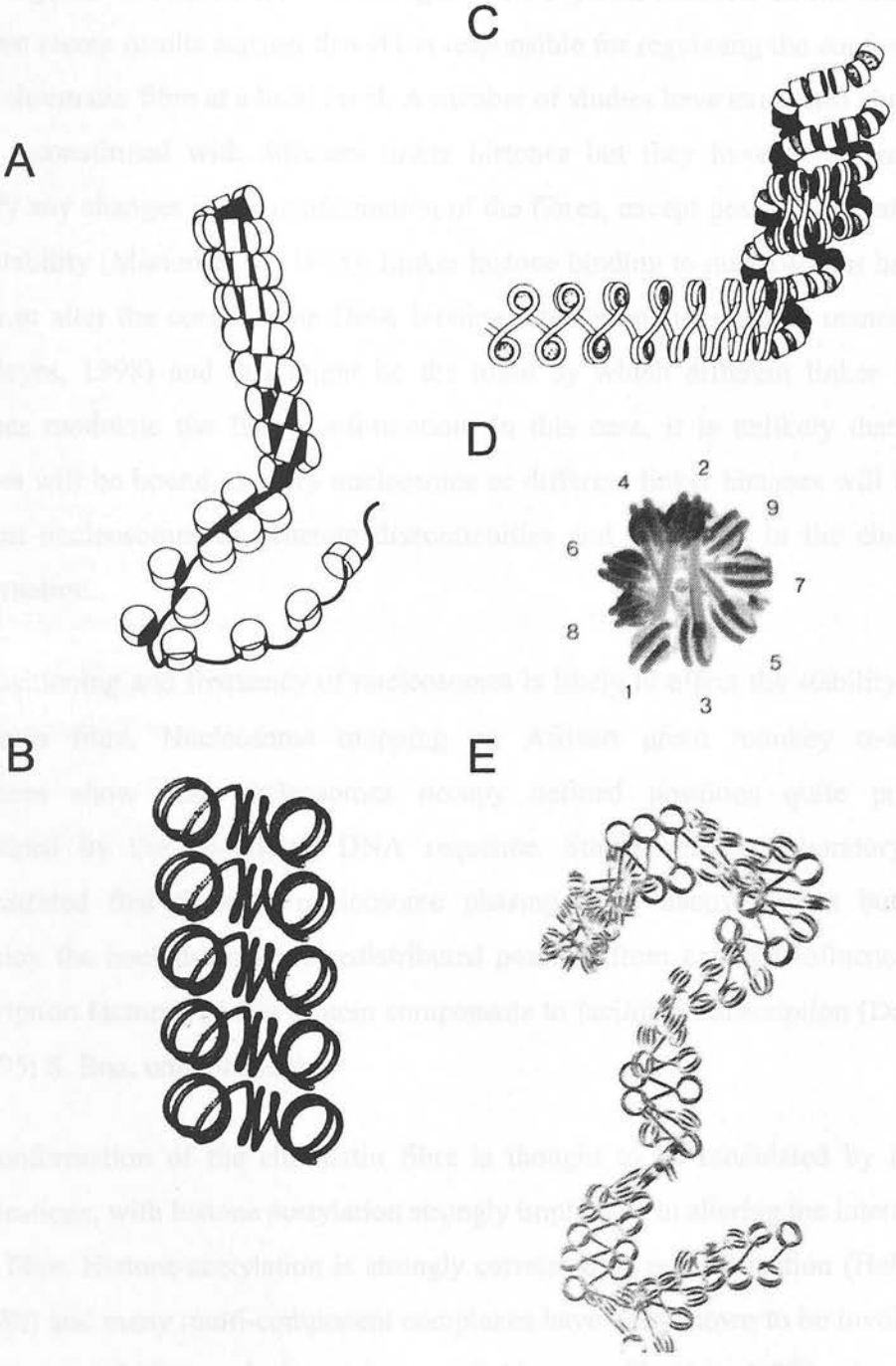


Figure 1.4 Models for the conformation of the 30nm chromatin fibre. (A) Solenoid (Thoma et al., 1979); (B) Solenoid variant (McGhee et al., 1983b); (C) Helical ribbon (Woodcock et al., 1984); (D) Crossed-linker (Williams et al., 1986); (E) Irregular fibre (Woodcock and Horowitz, 1995).

1.1.11 Modulating the conformation of the 30nm fibre

Originally, linker histone H1 was thought to be a global modifier of the chromatin fibre but recent results suggest that H1 is responsible for regulating the conformation of the chromatin fibre at a local level. A number of studies have examined chromatin fibres reconstituted with different linker histones but they have been unable to identify any changes in the conformation of the fibres, except possibly alterations in fibre stability (Marion et al., 1985). Linker histone binding to nucleosomes has been shown to alter the core-histone DNA binding interactions in a global manner (Lee and Hayes, 1998) and this might be the route by which different linker histone subtypes modulate the fibre conformation. In this case, it is unlikely that linker histones will be bound to every nucleosome or different linker histones will bind to adjacent nucleosomes to generate discontinuities and variations in the chromatin conformation.

The positioning and frequency of nucleosomes is likely to affect the stability of the chromatin fibre. Nucleosome mapping on African green monkey α-satellite sequences show that nucleosomes occupy defined positions quite probably determined by the underlying DNA sequence. Studies in our laboratory have demonstrated that there is nucleosome phasing over inactive genes but upon activation the nucleosomes are redistributed possibly from external influences like transcription factors or other protein components to facilitate transcription (Davey et al., 1995; S. Boa, unpublished).

The conformation of the chromatin fibre is thought to be modulated by histone modifications, with histone acetylation strongly implicated in altering the interactions of the fibre. Histone acetylation is strongly correlated to gene activation (Hebbes et al., 1988) and many multi-component complexes have been shown to be involved in both the acetylation and deacetylation of histones (Section 1.5.2). Acetylated chromatin is more sensitive to nucleases (Simpson, 1978b; Krajewski and Becker, 1998) and facilitates the binding of transcription factors to DNA in chromatin (Steger et al., 1998) suggesting that acetylation alters the conformation of the chromatin fibre.

For many years histone acetylation has been thought to modulate the conformation of the chromatin fibre, however sedimentation and electric dichroism studies have been unable to show that the conformation of acetylated fibres is any different to unacetylated fibres (McGhee et al., 1983a). This observation is corroborated by electron microscopy imaging, although at very high salt concentrations (100mM), control chromatin shows further compaction of the 30nm fibre, while hyperacetylated chromatin fails to undergo this final compaction step (Annunziato et al., 1988). The helical repeat of nucleosome bound DNA is unchanged by histone acetylation, and the extent of DNA-protein contacts remains unaltered (Bauer et al., 1994) although acetylation can release negative supercoils previously constrained by nucleosomes (Norton et al., 1989). In addition, the core histone N-termini which normally interact with DNA maintain this contact even when acetylated (Mutskov et al., 1998). *In vitro* reconstitution of DNA with acetylated histones compared to non-acetylated histones in a dinucleosome template facilitates transcription, but this effect is negated by the assembly of linker histones into chromatin (Ura et al., 1997). In the absence of linker histones the higher order folding of a 12-mer nucleosomal array is disrupted by a high level of core histone acetylation (Garcia-Ramirez et al., 1995) which concomitantly enhances transcription of the nucleosomal array by RNA polymerase III (Tse et al., 1998).

1.2 Cellular chromatin environment

Cells package their DNA so that it can be conveniently organised in the nucleus. In *Schizosacharomyces pombe,* the interphase chromatin fibre is packaged a further two-fold to form the metaphase fibre (Umesono et al., 1983) whilst in higher eukaryotes it has been calculated as approximately nine-fold (Rattner and Lin, 1985). It is not known how chromatin is packaged in a metaphase chromosome due to the inaccessibility of the cell nucleus for biochemical manipulation and microscopic analysis so many of the studies undertaken to date have only been able to elude the structures formed. However, low-angle X-ray scattering studies of isolated metaphase chromosomes show a series of diffraction features characteristic of 30nm chromatin fibres suggesting that metaphase chromosomes are produced by further folding of the basic 30nm fibre (Paulson and Langmore, 1983; Langmore and Paulson, 1983).

1.2.1 Nuclear chromatin organisation

Individual chromosomes can only be seen from early prophase to late telophase of the cell cycle. Consequently, much of what we know about nuclear organisation has come from studying mitotic cells. This only gives us part of the picture especially as eukaryotic DNA is normally transcribed and replicated in the context of an interphase nucleus. In contrast to the original idea that chromatin from mitotic chromosomes was randomly and diffusely dispersed throughout the interphase nucleus it is found that chromatin is non-randomly organised in a cell-specific manner with heterochromatin being localised to the interior of chromosome territories whilst active and inactive genes are located at the periphery (Kurz et al., 1996; Csink and Henikoff, 1998). Double labelling of cells at early and late S-phase has also allowed sub-chromosomal foci to be identified, and shown to be composed of either late or early replicating DNA (Zink et al., 1998). This approach has also been used in living cells to identify the dynamic movements of chromosomes. Chromosome axes follow simple pathways through their territories during G_2 phase and late replicating regions maintain their relative positions as prophase chromosomes form (Manders et al., 1999).

1.2.2 Euchromatin and heterochromatin

When interphase cells are examined microscopically using stains or fluorescent intercalating dyes (Figure 1.5) it is possible to distinguish two types of chromatin, heterochromatin and euchromatin. Facultative heterochromatin describes euchromatin which assumes the characteristics of heterochromatin in a developmentally controlled manner, for example the inactive copy of the female X-chromosome. Heterochromatin is condensed and appears brightly staining during interphase. It often contains repetitive genes and is generally transcriptionally silent, whereas euchromatin decondenses during interphase and appears more extended to facilitate the transcription of genes coding for cellular proteins. At metaphase, all chromatin is uniformly heterochromatic, therefore specific molecular signals that can direct appropriate local decondensation must be in place at the end of mitosis for the cell to enter back into a stable interphase state. Euchromatin and heterochromatin exhibit different acetylation patterns. Polyclonal antibodies against specific acetylation sites in histone H4 (Turner and Fellows, 1989) have been used to show that potentially active euchromatin can be modified at all the H4 acetylable lysines, whereas H4 in heterochromatin is hypoacetylated (O'Neill and Turner, 1995). The acetylation of coding regions in HL-60 cells is not altered upon differentiation although there is a transient increase in the acetylation of centromeric heterochromatin (O'Neill and Turner, 1995).

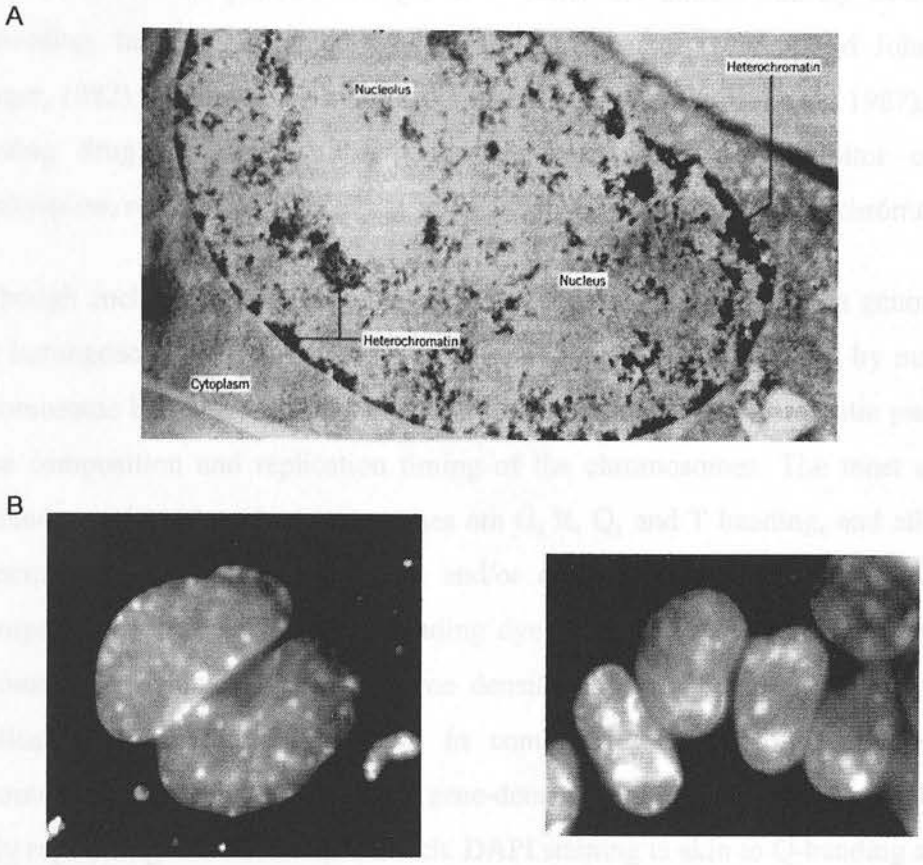


Figure 1.5 Micrographs showing heterochromatin and euchromatin. (A) Feulgen staining of a section through an interphase cell. The heterochromatin is visible as darkly staining condensed clumps (Picture from Genes III, B. Lewin). (B) DAPI staining of interphase nuclei. Heterochromatin is visible as brightly staining regions whilst the euchromatin is diffusely distributed.

For many years metaphase chromosomes have been extensively studied by chromosome banding. Constitutive heterochromatin can be specifically stained by heat treatment (C banding) and in man it is found in and around the centromere and the telomeric ends of the Y chromosome, whereas in other eukaryotes it is found in both telomeric and pericentric regions. C bands are almost entirely occupied by noncoding, tandemly repeated satellite DNA sequences (Miklos and John, 1979; Singer, 1982). Treatment of living cells with distamycin (Radic et al., 1987), an AT-binding drug, or 5-azacytidine (Joseph et al., 1989), an inhibitor of DNA methylation, results in a marked decondensation of centromeric heterochromatin.

Although euchromatin contains the majority of genes in a mammalian genome, it is not homogenous in its response to staining and can be differentiated by metaphase chromosome banding. Differential staining methods reflect the chromatin packaging, base composition and replication timing of the chromosomes. The most common methods used to identify chromosomes are G, R, Q, and T banding, and all use the principle of chromatin denaturation and/or enzymatic digestion followed by the incorporation of a specific DNA-binding dye (Figure 1.6). G bands correspond to chromomeres which have a low gene density, low number of CpG islands, late replicating and relatively AT-rich. In contrast, R bands correspond to inter-chromomere DNA which has a high gene-density, higher number of CpG islands, is early replicating and relatively GC-rich. DAPI staining is akin to Q-banding although it is less distinct that quinacrine. The Q-positive bands are also G-positive indicative of being AT-rich. Banding techniques provide an important tool for locating gene rich regions and karyotyping chromosomes but unfortunately they tell us very little about the underlying chromatin structure.

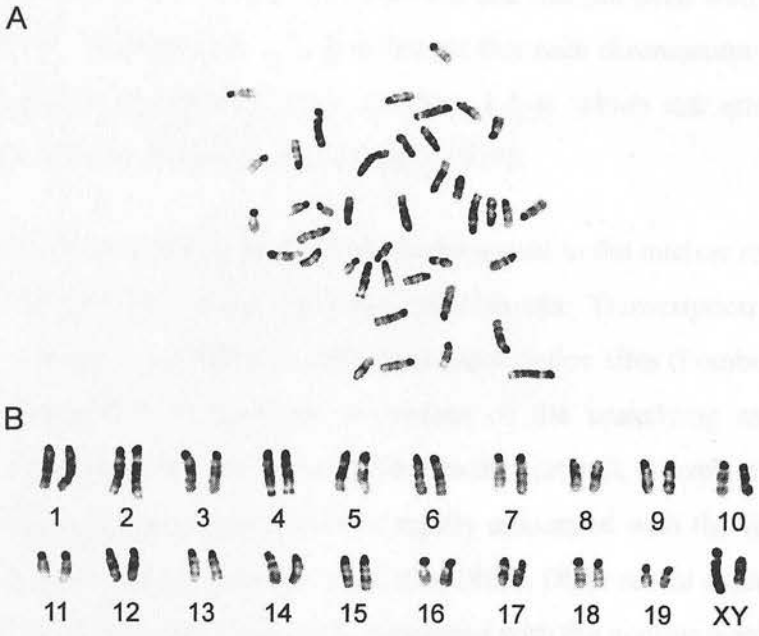


Figure 1.6 Enhanced DAPI banding. (A) Mouse embryonic stem cell metaphase spread counter-stained with DAPI. The image was collected with a CCD camera and electronically enhanced. (B) Karyotype analysis of (A). (Picture taken from Bickmore and Craig, 1997).

1.2.3 Nuclear matrix

Cellular chromatin does not float freely in the nucleoplasm; it is associated with the nuclear membrane and other components of the nuclear scaffold in a regulated manner such that the matrix is an important factor in the control of gene expression (Strouboulis and Wolffe, 1996). The nuclear matrix can be operationally defined as the nuclear structure that remains after the salt extraction of nuclease-treated nuclei (i.e. histone depleted). It consists of peripheral lamina-pore complexes and an internal filamentous ribonucleoprotein network that has not been well characterised (Berezney et al., 1995) (Figure 1.7). It is thought that each chromosome is subdivided into topologically constrained loops (Section 1.2.4) which are attached to this proteinaceous matrix (Marsden and Laemmli, 1979).

In addition to the physical attachment of chromosomes to the nuclear matrix, there is much evidence for functional chromatin attachments. Transcription by all three nuclear polymerases is localised to dedicated transcription sites (Pombo et al., 1999) which are thought to be localised to regions of the underlying nuclear matrix providing attachments for the chromatin fibre to this network (Hozak et al., 1994). In addition, transcriptionally active DNA is tightly associated with the nuclear matrix whilst inactive loci are not (Jackson and Cook, 1985). Other recent evidence suggests that splicing complexes are also closely associated with the nuclear matrix (Misteli et al., 1997). Specific transcription-related factors are transported in interchromatin granules from perichromatin fibrils to sites of active DNA along components of the matrix thus acting as functional associations between the chromatin and the underlying nuclear matrix (Misteli and Spector, 1998). The SWI/SNF chromatin remodelling complex (Section 1.5.2) is tightly associated with 11 polypeptides. In yeast, two of these, SWP59 and SWP61 are encoded by the ARP9 and ARP7 genes, respectively, which code for members of the actin-related protein family (Peterson et al., 1998). In addition, actin has been found associated with the human BRG1 (SWI/SNF homolog) complex (Zhou et al., 1998) suggesting that this complex might interact with some component of the nuclear matrix.

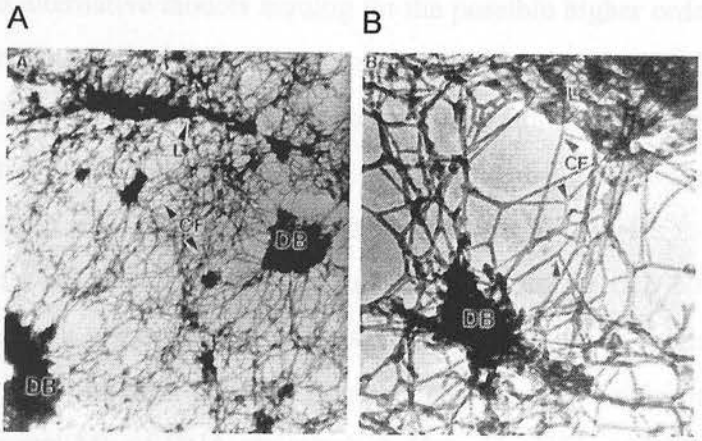


Figure 1.7 Ultrastructure of nuclear matrix core filaments. Micrographs at low (A) and high (B) magnification show the complex interactions between the nuclear matrix components. (CF) Core filaments, (DB) Dense bodies, (L) Nuclear lamina. (Picture taken from Penman, 1995).

1.2.4 Cellular higher order chromatin fibre

The conformation of metaphase chromosomes has been studied by a combination of microscopic analysis of salt or detergent extracted nuclei. One of the earliest models for the organisation of a metaphase chromosome was the radial loop model (Paulson and Laemmli, 1977; Marsden and Laemmli, 1979; Rattner and Lin, 1985; Boy de la Tour and Laemmli, 1988). In this model the 30nm chromatin fibres are arranged into loops by regular attachments to a proteinaceous scaffold structure which forms the backbone of the chromosome (Figure 1.8). Other studies using different techniques have proposed alternative models arguing for the possible higher order arrangements of 30nm fibres into large scale chromatin structures (Belmont et al., 1999) such as the 100- to 130nm chromonema fibre (Belmont and Bruce, 1994) or a 240nm fibre (Manuelidis and Chen, 1990). Filipski et al. (1990b) using whole cell micrococcal nuclease digestion showed that the cellular chromatin was digested to 50 and 300 kb DNA fragments suggesting an arrangement of the chromatin into 50 kb loops with six loops forming a higher 300 kb organisation. The random walk/giant loop model (Sachs et al., 1995) predicts a highly flexible back-bone to which giant loops, each comprising several Mb of DNA are attached. However, this model does not predict an ordered organisation of a chromosome territory.

The strongest evidence does suggest that the 30nm fibre is condensed by the formation of an axial loop structure, and by the further compaction of these loops much like a bottle brush to create the metaphase chromosome (Pienta and Coffey, 1984a; Callan, 1986) (Figure 1.9). On average, each chromosome would have about 1800 loops arranged radially, possibly in a solenoidal fashion. Pienta & Coffey (1984a) further subdivided this into 100 units containing 18 loops per turn called mini-bands. In this model, the stability of the metaphase chromosome could be increased by the different levels of folding having alternate handedness much like the winding of a rope.

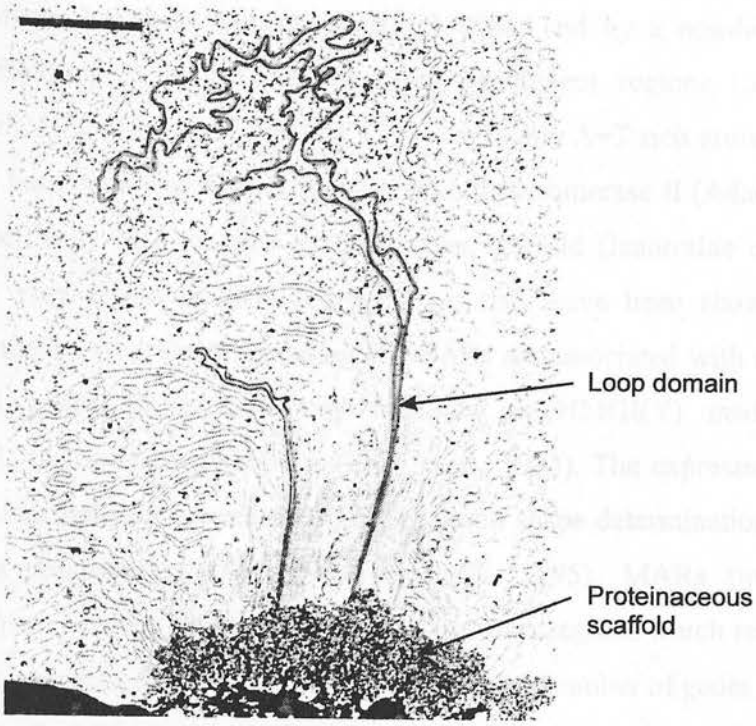


Figure 1.8 Chromosomal DNA is attached to a protein scaffold. Electron micrograph of a histone depleted chromosome showing the DNA attachments (Picture taken from Paulson and Laemmli, 1977).

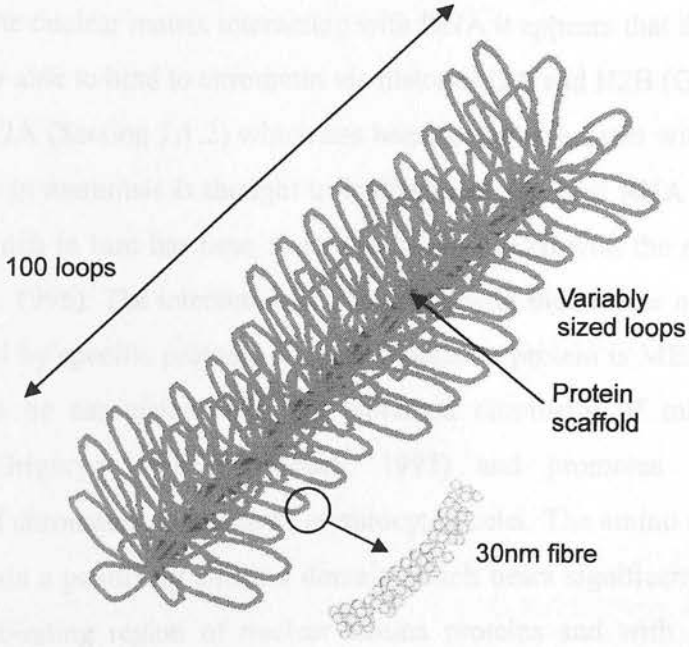


Figure 1.9 Radial loop model for the formation of the metaphase chromosome fibre. Chromosomal DNA is looped into variably sized domains with an average size of 80 kb. These domains are grouped into mini-bands of about 18 loops organised radially around the protein scaffold with approximately 100 of these minibands arranged along the length of the chromosome, possibly in a solenoidal fashion.

The axial loops of chromosomes are constrained at either end by a non-histone proteinaceous scaffolding structure. These scaffold attachment regions (SARs) (Strick and Laemmli, 1995; Hart and Laemmli, 1998) are highly A+T rich sequences of several hundred base pairs which are associated with topoisomerase II (Adachi et al., 1989) and specifically bind *in vitro* to the nuclear scaffold (Izaurralde et al., 1989). SARs may play a role in gene expression as they have been shown to selectively bind a number of proteins including H1. SARs are associated with many highly transcribed genes where they may facilitate an HMGI(Y) mediated displacement of H1, facilitating transcription (Zhao et al., 1993). The expression of synthetic SARs shows that they are required for chromatid shape determination but not for chromatin condensation (Strick and Laemmli, 1995). MARs (matrix attachment regions) are possibly related to SARs and are DNA regions which remain attached to the nuclear matrix after extracting the chromatin. A number of genes have MARs, such as the β-globin locus, c-myc promoter and DHFR gene (Boulikas, 1995).

In addition to the nuclear matrix interacting with DNA it appears that the tail domain of lamin Dm_0 is able to bind to chromatin via histone H2A and H2B (Goldberg et al., 1999). MacroH2A (Section 1.1.2) which has been found associated with the inactive X-chromosome in mammals is thought to interact with the Xist RNA (Csankovszki et al., 1999) which in turn has been shown to be associated with the nuclear matrix (Clemson et al., 1996). The interaction of chromatin with the nuclear matrix is likely to be modulated by specific proteins. One such putative protein is MENT which has been shown to be associated with the repressed chromatin of mature chicken erythrocytes (Grigoryev and Woodcock, 1993) and promotes the *in vitro* condensation of chromatin in immature erythrocyte nuclei. The amino acid sequence of MENT reveals a positively charged domain which bears significant similarity to the chromatin binding region of nuclear lamina proteins and with the A+T-rich binding motif, which may account for the targeting of MENT to peripheral heterochromatin (Grigoryev et al., 1999).

1.2.5 Junk DNA

The non-coding DNA, often referred to as junk DNA (Zuckerkandl, 1997), constitutes more than 90% of the mammalian genome and includes a variety of sequence classes such as satellite DNA, long interspersed repeated elements and smaller DNA sequence motifs. These DNA sequences were rarely considered to have a function but increasingly they are thought to play a role within selected chromosome domains and participate in more complex levels of chromatin folding, or the indexing of different genetic compartments for orderly transcription and replication.

1.2.6 Position-effect variegation

Heterochromatin is normally considered to be a repressive environment (Hennig, 1999). Heterochromatin protein 1 (HP1) was originally described in *Drosophila* (James and Elgin, 1986), and is found to be responsible for the phosphorylation-dependent packaging of heterochromatin (Eissenberg et al., 1994). It is conserved across a number of species including mammals. The *Tetrahymena* homolog is found to be responsible for silencing of genes in the macronuclei (Huang et al., 1998; Huang et al., 1999) and M31, a mouse homolog is localised to constitutive heterochromatin (Wreggett et al., 1994) whereas M32, another HP1 homolog, is localised to euchromatin (Horsley et al., 1996). The PEV-modifier Suv39h1/SUV39H1 encodes a centromere associated protein which complexes with M31 indicating that, as for *Drosophila* (Cleard et al., 1997), there is a mammalian SU(VAR) complex which may co-operate in building the genomic silencing associated with heterochromatin (Aagaard et al., 1999).

In addition to heterochromatin being coated with specific proteins it is hypoacetylated which maybe required for its condensed architecture. In *Drosophila* and yeast, heterochromatin is acetylated at H4 K12 (Turner et al., 1992; Braunstein et al., 1996) whilst human chromatin is acetylated at K12 and/or K16. Unlike euchromatin, assembled heterochromatin excludes many enzymes and if acetyltransferases and deacetylases are among these the remaining acetylated sites

may remain preferentially acetylated. Tumbar et al. (1999) showed that a 90Mbp heterochromatin chromosome could be decondensed by targetting a VP16-lac repressor fusion to chromosome regions containing lac operator sequences, which caused a concomitant increase in transcription, localised histone hyperacetylation, and recruitment of histone acetyltransferases. These results indicate a functional link between the recruitment of transcriptional machinery and changes in large-scale chromatin structure, possibly suggesting a reason for finding housekeeping genes localised in clusters. In addition to heterochromatin-specific proteins there are some factors which are localised to euchromatin, for example M32 (Horsley et al., 1996) or Tif1α have been found to be tightly associated with highly accessible euchromatic regions of the genome (Remboutsika et al., 1999).

When euchromatic genes are brought into juxtaposition with heterochromatin by chromosomal rearrangement or transposition, they can exhibit position effect variegation (PEV). This effect is manifested by a silencing of the gene in a subset of cells in which it is not normally expressed and because of this PEV is now being associated with certain human diseases (Kleinjan and Van Heyningen, 1998). A number of PEV modifiers have been identified in genetic screens in *Drosophila* and yeast and many of these proteins are associated with heterochromatin. Localisation of heterochromatin in a euchromatic gene also causes PEV, and induces the aberrant association of the gene with heterochromatin (Csink and Henikoff, 1996). Other data suggest that in *Drosophila*, both nuclear organisation and local chromatin structure play a role in both centromeric (Dernburg et al., 1996a; Dernburg et al., 1996b) and telomeric PEV (Cryderman et al., 1999). In *Drosophila* it has been shown using bromodeoxyuridine incorporation in combination with fluorescence in situ hybridisation that there are heterochromatic interactions between distal and proximal parts of the chromosome arms at specific times at the cell cycle suggesting a mechanism that links gene regulation via nuclear positioning to the cell cycle (Csink and Henikoff, 1998).

Studies in mammalian cells have extended these observations by showing that several transcriptionally inactive genes localise to centromeric heterochromatin in the

nucleus of cycling but not quiescent primary B lymphocytes. These results indicate that the spatial organisation of genes in cycling and non-cycling lymphocytes is different and that locus repositioning may be a feature of heritable gene silencing (Brown et al., 1999). In addition these inactive heterochromatin foci are associated with the transcription activator, Ikaros (Brown et al., 1997a), which upon T cell activation recruits NuRD, a histone deacetylase, (Section 1.5.2) to regions of heterochromatin such that it maybe able to modify the chromatin structure of specific genes (Kim et al., 1999). This corroborates with recent evidence which shows that the NuRD complex interacts with MBD2, a methylated DNA binding protein, and may play a role in silencing methylated regions such as heterochromatin (Zhang et al., 1999). Ikaros has also been shown to interact with the related protein Helios which is found in limiting quantities and localises extensively to heterochromatin suggesting it may be a regulator of Ikaros (Hahm et al., 1998).

1.3 Centromeres

The primary morphological constriction of chromosomes is the centromere (Choo, 1997; Craig et al., 1999). Centromeres have multiple roles during mitosis, and may also function during interphase (Pluta et al., 1995). They are responsible for directing the formation of the kinetochore, a button-like structure at the chromosomal surface which attaches to the spindle microtubules and regulates chromosome movements in mitosis. The lateral borders of the centromere are the last points at which the sister chromatids remain paired prior to their disjunction at anaphase, and presumably they contain the machinery necessary to effect this separation. The centromere is also involved in cell cycle checkpoint control and in most cells, centromeres that have not achieved a stable bipolar orientation on the spindle delay the onset of metaphase-anaphase transition. In most situations, the position of the centromere on a given chromosome is fixed, suggesting that its location is defined by specific DNA sequences, structures or epigenetic effects.

1.3.1 Centromeric satellite arrays

Most centromeres are composed of a remarkably complex array of repeat structures (Figure 1.10). However, studies on neocentromeres suggest that repetitive DNA is not a prerequisite for centromere activity (Eichler, 1999). The DNA sequence at centromeres may vary between species but the molecular architecture is generally the same: blocks of AT-rich tandemly repeated DNA sequences bracketed by clusters of various classes of retrotransposons. The best characterised of the centromeric satellite sequences is α-satellite DNA, a primate-specific family of satellite DNA based on a fundamental repeat of 171 base pairs (bp) (Fittler and Zachau, 1979) which comprises up to 5% of the total genomic DNA. It was originally identified as a tandemly repeated 340 or 680 bp sequence that hybridised to the centromeric regions of all human chromosomes, though the Y chromosome α-satellite is smaller than the arrays found on the other chromosomes. Mouse minor satellite is thought to be functionally equivalent to human α-satellite and is based on a tandem repeat length of 120 bp (Wong and Rattner, 1988) and it is related to a portion of the 234 bp mouse major satellite repeat (Horvath and Horz, 1981). In mouse satellite DNA, seven of the eight CpG sites are methylated (Manuelidis, 1981) and its histones are hypoacetylated (Pashev et al., 1983). Nucleosome positioning sites have been mapped on the 171 base pair α-satellite DNA in African green monkey cells and are shown to occupy eight strictly defined positions (Zhang et al., 1983). Introduction of human α-satellite DNA into ectopic chromosome sites in African green monkey cells is associated with bridges between the separating sets of chromatids at anaphase and an increased number of lagging chromosomes at metaphase (Haaf et al., 1992). This suggests that α-satellite DNA provides the primary sequence information for centromere binding and for at least some functional aspects of a mammalian centromere. However, more recent studies, have indicated that the effect of ectopic α-satellite on chromosome segregation may be due to delayed sister chromatid separation caused by the repetitive nature of the DNA and does not reflect the formation of centromeres (Warburton and Cooke, 1997).

The only centromeric satellite sequences found in mouse are the major and minor satellites. In contrast a variety of different sequences have been described in humans. The classical satellites, satellites 1-3, were among the first tandemly repeated DNA characterised in the human genome. These satellites are made up of short repeats that are AT rich. The monomeric repeating unit for satellite 1 DNA is 42 bp whilst satellite 2 and 3 are based on a 5 bp monomeric unit (Prosser et al., 1986). Satellite 1 is found on chromosome 3, 4 and all the acrocentric chromosomes. On chromosomes 3 and 4 it maps to the proximal q arm side of the centromeric α-satellite (Tagarro et al., 1994). Satellites 2 and 3 are present in various amounts at the pericentromeric regions of most human chromosomes.

Human β-satellite is another class of centromeric satellite. It exists as a tandem array of diverged 68 bp monomer repeats up to 1 Mb in length (Waye and Willard, 1989) that replicate late in S phase of the cell cycle (Ten Hagen and Cohen, 1993). They are specifically found on the acrocentric chromosomes (13,14,15,21,22) and to a lesser extent on other chromosomes. γ-Satellite is the newest family of repeats to be described for chromosome 8 and X. They are composed of a tandem monomer of ~220 bp which has been localised to the lateral sides of the primary constriction. Two other families of repeat have been described: the 48 bp repeat and the ATRS (AT rich sequence). The 48 bp repeat is only found on chromosomes 21 and 22 close to the α-satellite, whilst the ATRS is an interspersed repeat found at or near centromeres.

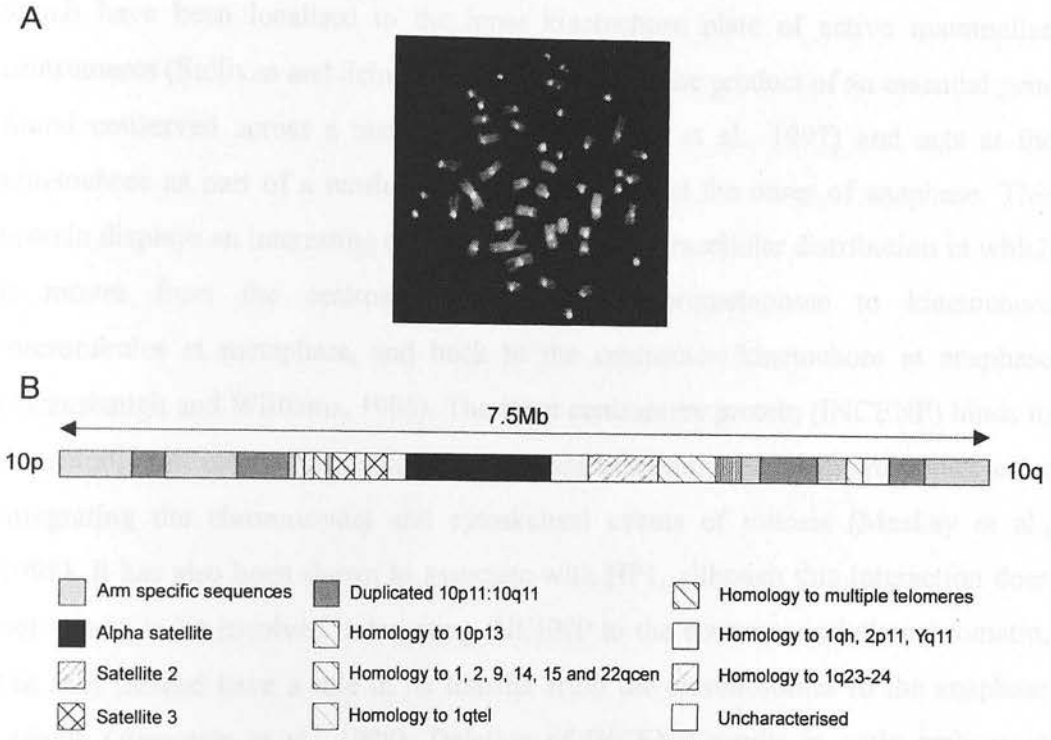


Figure 1.10 DNA sequence organisation around human centromeres. (A) Human metaphase chromosome spread counter stained with SYTOX green and hybridised with an α-satellite all chromosome probe (Picture from Molecular Probes). (B) Sequence organisation across the 7.5Mb centromere of human chromosome 10 (Picture taken from Jackson et al., 1999).

1.3.2 Centromere proteins

A variety of centromere-associated proteins have been identified, many using serum from CREST scleroderma patients (Moroi et al., 1980), which bind to the kinetochore region of centromeres. Of these, CENP-A (Sullivan et al., 1994), CENP-B (Pluta et al., 1992) and pJα (Gaff et al., 1994; Romanova et al., 1996) are specifically associated with α-satellite DNA and CENP-G is associated with the α-1 satellite DNA subfamily (He et al., 1998). CENP-C and CENP-E are key proteins which have been localised to the inner kinetochore plate of active mammalian centromeres (Sullivan and Schwartz, 1995). ZW10 is the product of an essential gene found conserved across a number of species (Starr et al., 1997) and acts at the kinetochore as part of a tension-sensing checkpoint at the onset of anaphase. This protein displays an interesting cell cycle-dependent intracellular distribution in which it moves from the centromere/kinetochore at prometaphase to kinetochore microtubules at metaphase, and back to the centromere/kinetochore at anaphase (Strausbaugh and Williams, 1996). The inner centromere protein (INCENP) binds to the spindle midzone at the metaphase/anaphase transition and maybe responsible for integrating the chromosomal and cytoskeletal events of mitosis (Mackay et al., 1998). It has also been shown to associate with HP1, although this interaction does not appear to be involved in targeting INCENP to the centromeric heterochromatin, but may instead have a role in its transfer from the chromosomes to the anaphase spindle (Ainsztein et al., 1998). Deletion of INCENP results in early embryonic lethality with a cellular phenotype which suggests a defect in the modulation in microtubule dynamics (Cutts et al., 1999).

CENP-A is a 17kD centromere-specific histone variant that is related to the carboxy-terminal domain of histone H3 and to CSE4, a component of yeast centromeres (Meluh et al., 1998). CENP-A is detected at all active centromeres (Warburton et al., 1997) as well as at neocentromeres although it is absent from the inactive centromeres seen on dicentric chromosomes. Immunolocalisation of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres (Warburton et al., 1997). Microinjection of antibodies against CENP-A

arrests cells in interphase but does not prevent mitosis (Figueroa et al., 1998) suggesting that CENP-A is involved in an essential interphase event occurring at the centromere before mitosis. This may include chromatin assembly at the kinetochore, co-ordinate with late replication of satellite DNA to form an active centromere.

CENP-B is a centromeric DNA-binding protein whose recognition site is found in many primate α-satellite and mouse minor satellite sequences. The protein is found localised to centromeres but it is not found at neocentromeres or on the Y chromosome. Satellite demethylation by growing cells in 5-azacytidine causes CENP-B to decorate a larger area of the centromere than normal, suggesting that it preferentially binds to sequences which are methylated (Mitchell et al., 1996). Current evidence suggests this protein is not required for normal chromosome segregation in mammalian cells as CENP-B null mice are mitotically and meiotically normal and only have lower body and testis weights (Kapoor et al., 1998; Perez-Castro et al., 1998; Hudson et al., 1998). CENP-B binding may have an accessory role, such as altering the centromeric chromatin structure to facilitate efficient and stable association of other centromere proteins within this region.

CENP-C is the only centromere protein other than CENP-A that shows clear homology to a yeast centromere protein (Lanini and McKeon, 1995; Brown, 1995). Its yeast homologue, MIF2p, binds to centromeric DNA (Meluh and Koshland, 1997) and is encoded by an essential gene which when mutated or overexpressed causes errors in chromosome segregation (Brown et al., 1993). Disruption of the mouse CENP-C gene has shown that it is essential for normal development (Fukagawa and Brown, 1997; Kalitsis et al., 1998a) with mitotic arrest and gross morphological degeneration becoming apparent as early as the morula stage of development. CENP-C delocalisation from the centromere in chicken DT40 cells causes disassembly of the centromere protein complex and blocks cells at the metaphase-anaphase junction. In contrast, over expression of CENP-C causes an increase in errors of chromosome segregation and inhibits the completion of mitosis (Fukagawa et al., 1999) but is not sufficient to induce formation of de novo functional centromeres.

HMGI(Y) has been shown to bind to multiple sites in mouse satellite repeats leading to the suggestion that it may be involved in heterochromatin condensation (Radic et al., 1992). In addition it has been proposed to play a role in organising nucleosome placement in centromeric satellite repeats (Strauss and Varshavsky, 1984). Also, MeCP-2 binds directly to centromeric DNA sequences of mammalian chromosomes (Lewis et al., 1992).

1.3.3 Neocentromeres

Neocentromeres can either be found naturally or be induced to occur by cellular damage. In *Drosophila* a neocentromere has been generated by irradiation and this acentric minichromosome binds the specific centromere protein ZW10. Analysis has shown that the sequences contained in this acentric mini-chromosomes were derived from the tip of the X-chromosome, which does not normally exhibit centromere activity (Williams et al., 1998). In addition, this sub-telomeric DNA does not display neocentromere function in the original X-chromosome context when separated from the rest of the X chromosome suggesting that neocentromere function requires an activation step. Possibly, centromere activity is normally repressed by lateral inhibition by the functional X centromere or perhaps the centromere function of the acentric mini-chromosome was activated by previous proximity to a functional centromere.

Functional human neocentromeres have been identified that possess no detectable α-satellite, satellite III, or CENP-B protein (Voullaire et al., 1993) and a neocentromere on human chromosome 3 has been found without detectable α-satellite DNA but which forms morphologically normal kinetochores (Wandall et al., 1998). Also, neocentromeres have been shown to stably pass through three human generations with no detrimental affects (Tyler-Smith et al., 1999). Sequence analysis of an 80kb human neocentromere (Barry et al., 1999) on chromosome 10 is devoid of human centromeric α-satellite DNA and the pericentric β- and γ- satellites, the ATRS and 48 base pair repeat DNA. In this case CENP-A, CENP-C, CENP-E could not be detected at the site of the canonic centromere but were present at the new constriction. It is currently unclear what occurs during the generation of a

neocentromere: is it stimulated by deletion of the normal centromere or does the formation of a new centromere induce the redundancy of the 'old' centromere?

The stability of certain dicentric chromosomes in humans seems to result from inactivation of one centromere, yielding a functionally monocentric chromosome. In a dicentric (X;15) translocation, CENP-C was never found at the inactive chromosome 15 centromere, although both centromeres carry α-satellite DNA (Page et al., 1995). Identification of centromeric antigens in dicentric Robertsonian translocations have shown that CENP-C and CENP-E are necessary components of functional centromeres and studies on chromosome 14 and 15 Robertsonian translocations suggest that there is an active centromere preference with the chromosome 14 centromere most often activated (Sullivan and Schwartz, 1995).

1.3.4 Mammalian artificial chromosomes

The generation of *S. cerevisiae* chromosomes (YACs) has proved relatively straightforward (Murray and Szostak, 1983). After identifying the centromere, telomere and a suitable replication origin it was possible to join them to construct a functional chromosome. In contrast it has been difficult to generate a mammalian artificial chromosome (MAC). Although telomere sequences have been identified the minimal components of a mammalian centromere are unclear (Masumoto et al., 1998; Willard, 1998). Construction of MACs with telomere, centromere and replication function have been approached by two methods. The top down strategy uses artificially induced chromosome truncations as a means to define a minimal chromosome that retains the mitotic properties of a normal chromosome. The 'build up' approach has focused on attempts to assemble MAC vectors containing functionally defined telomere repeats together with candidate centromere sequences.

Recently these approaches have begun to yield rewarding results. By combining long synthetic arrays of α-satellite DNA with telomeric DNA and genomic DNA in human HT1080 cells it has been possible to generate a stable chromosome (Harrington et al., 1997). However, the size of this chromosomes is estimated to be 6-10 Mb which approximates to one-fifth to one-tenth the size of endogenous human

chromosome so it is unlikely that these centromeric sequences are minimal. By inserting telomere sequences and specific α21-I alphoid DNA, which contains CENP-B boxes, into a YAC it is possible to form a stable chromosome in the absence of selection (Masumoto et al., 1998) although once again these constructs are 1-5 Mb in size. A 2.4 Mb human X centromere-based minichromosome has also been developed by targeted telomere-associated chromosome fragmentation in DT40 cells (Mills et al., 1999).

By understanding the key protein components and underlying sequence requirements involved in forming a functional centromere it may be possible to develop a smaller construct. To date all of the artificial chromosomes made have used α-satellite; however if the key feature of the centromere is more structural (e.g. rigidity of the DNA sequence) it might be more efficient to use the type of sequence found in neo-centromeres, or a totally artificial sequence which has been designed to have a particular structural conformation.

1.3.5 Centromere determinants

Emerging repetitive structures at centromeric sites may be an important by-product of a functional centromere which ensures that the site remains an evolutionarily favoured position in subsequent meiotic and mitotic lineages. This might be through a system of epigenetic marking. Suitable candidates would be proteins which remain at the centromere throughout the cell cycle, namely CENP-A, CENP-C, and pJα. This in itself may not be sufficient to maintain the centromere and over time there could be a tendency for centromere drift to occur, so the addition of specific DNA sequence may be responsible for further stabilising the system.

Another candidate for marking the centromere may simply be due to it being the last sequence to replicate (DuPraw, 1968; Csink and Henikoff, 1998). Possibly, the accumulation of repetitive DNA further retards replication, effectively fixing such sites in the genome to function more competently as a centromere. In other words, centromere function is not a consequence of repetitive structure, but repetitive structure is a consequence of enhancing centromere function. It has been found that

the transcription start site of active genes are marked during mitosis by a subset of proteins that distort the DNA conformation (Michelotti et al., 1997), and to some extent proteins at the centromere such as CENP-A may play a similar role in altering the conformation of the underlying DNA to make it function more efficiently as a centromeric region.

1.4 Telomeres

Bacterial genomes, plasmids, bacteriophages and mitochondrial DNAs tend to be circular. In contrast eukaryotic cellular chromosomes are normally linear. A major disadvantage of this organisation is the presence of free DNA ends in the eukaryotic nucleus. Telomeres (Kipling, 1995; Pardue and DeBaryshe, 1999) were first identified as distinctive structures at the natural ends of chromosomes by cytological and genetic studies in *Drosophila* (Muller, 1938). Telomeres play important roles in maintaining chromosome stability, complete DNA replication, correct chromosome segregation and positioning of chromosomes within the nucleus. Due to non-conservative replication of DNA, organisms must possess mechanisms to maintain telomere length, and without such mechanisms telomeric sequences are lost during each round of replication (Lingner and Cech, 1998). The telomeres of mammals are constructed of closely spaced nucleosomes, in contrast to the telomeres of lower eukaryotes which show no evidence of a nucleosomal structure (Makarov et al., 1993). They also have a bipartite structure with unusual chromatin near the telomere terminus and a more canonical nucleosome organisation in the proximal part of the telomere (Sutcliffe et al., 1994). The lengths of telomeres gradually shorten as fibroblasts are cultured (Harley et al., 1990) and studies in humans have shown that telomere length in different tissues decreases with increasing age (Butler et al., 1998). An analysis of telomere lengths in cloned sheep suggest that the lengths of the telomeres are slowly reduced during the lifetime of the animal and can not be reset by passing through a new oocyte (Shiels et al., 1999).

Telomerase is a specialised reverse transcriptase that synthesises telomeric repeats de novo (Nugent and Lundblad, 1998). The catalytic subunit of telomerase (TERT) is encoded by est2 in *S. cerevisiae* and it is a reverse transcriptase that is closely related

to the reverse transcriptases of non-long terminal repeat retrotransposons (Nakamura et al., 1997). Most metazoan telomeres contain short G-rich tracks which are laid down by telomerase using an RNA primer as a template (Greider, 1996). Mice deficient for the telomerase RNA and lacking telomerase activity can only be bred for approximately six generations due to increased male and female sterility (Blasco et al., 1997; Lee et al., 1998; Rudolph et al., 1999). The viability of these mice can be further reduced to four generations if they are bred into a mouse background which has shorter telomeres (Herrera et al., 1999b). A variety of disease states associated with mouse ageing appear on this background including sterility, splenic atrophy, reduced proliferative capacity of B and T cells, abnormal haematology and atrophy of the small intestine.

Despite the importance of telomere maintenance, it is clear that telomeres do not simply function as a buffer zone that prevents loss of essential sequences and end-to-end chromosome fusions, but more importantly that the telomeric complex allows cells to distinguish random DNA breaks from natural chromosome ends. Whereas broken chromosomes activate DNA damage checkpoints (Sandell and Zakian, 1993) and are repaired, telomeres are not detected as DNA ends. The molecular mechanism for this capping function has been proposed to depend either on a specific DNA structure in the most terminal single-stranded portion of the telomere or require proteins bound to the telomere end. Strong evidence suggests that the identity of telomeres may be a combination of structure and proteins bound to the ends of the chromosomes. TRF1 (telomeric-repeat binding factor) and TRF2 are two human proteins that bind to the human telomeric sequence. TRF1 is a negative regulator of telomere length maintenance (van Steensel and de Lange, 1997) whilst TRF2 binds to telomeres and blocks end-to-end fusions (van Steensel et al., 1998). TRFF2 elicits an architectural change in the conformation of the telomeric DNA (Figure 1.11), in effect a molecular knot which provides a general mechanism for the protection and replication of telomeres. (Griffith et al., 1999).

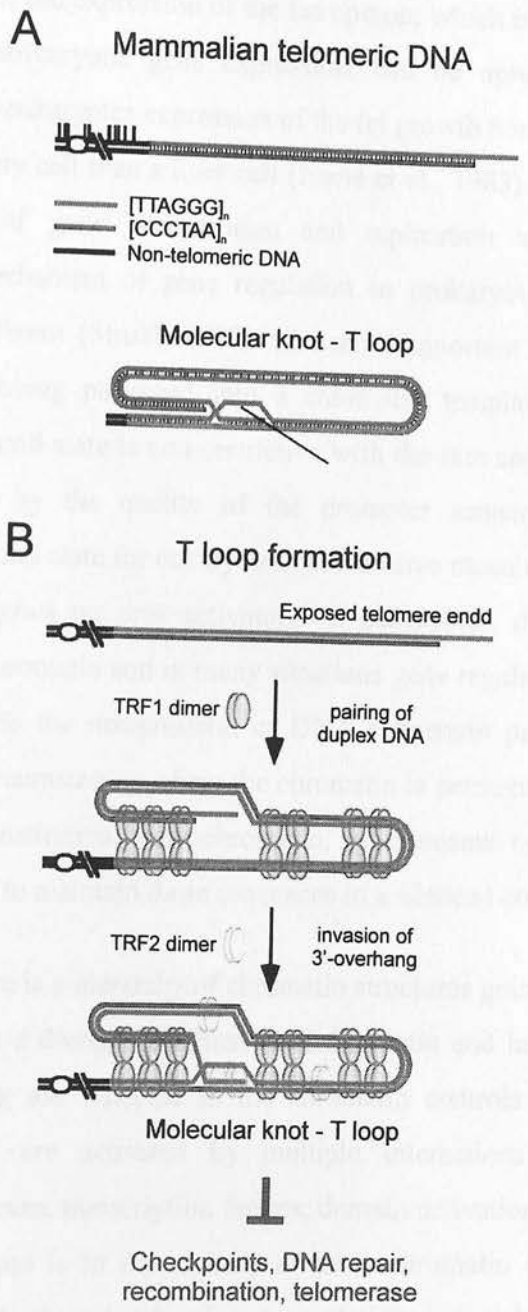


Figure 1.11 Model for the stabilising T-loop found at the end of telomeres. T-loops are proposed to mask telomere termini from cellular activities that can act on DNA ends (A) Proposed conformation of telomeric DNA. (B) Interaction of telomere proteins TRF1 and TRF2 are proposed to facilitate the formation of the T-loop. (Picture taken from Griffith et al., 1999).

1.5 Chromatin and the regulation of gene transcription

In *Escherichia coli*, the expression of the lac operon, which is often considered to be a paradigm for prokaryotic gene expression, can be upregulated 10^4 times. In contrast, in higher eukaryotes expression of the rat growth hormone gene is 10^9 times greater in a pituitary cell than a liver cell (Ivarie et al., 1983). Although many of the basic principles of gene transcription and replication are universal between organisms the mechanism of gene regulation in prokaryotes and eukaryotes are fundamentally different (Struhl, 1999). The most important difference arises from eukaryotic DNA being packaged into a chromatin template. In prokaryotes the transcriptional ground state is non-restrictive with the rate and level of transcription being determined by the quality of the promoter sequences. In contrast, the transcriptional ground state for eukaryotes is restrictive meaning that transcription of most eukaryotic genes requires activators. In eukaryotes, the global repressor is considered to be chromatin and in many situations gene regulation in eukaryotes can be considered to be the derepression of DNA sequences packaged in chromatin. However, under circumstances where the chromatin is permanently inactivated, as in facultative and constitutive heterochromatin, mechanisms operate in conjunction with the chromatin to maintain these sequences in a silenced condition.

Within the cell there is a hierarchy of chromatin structures going from a nucleosome, to a 30nm fibre, to a domain and then to euchromatin and heterochromatin. These levels of packaging are reflected in the chromatin controls which regulate gene expression. Genes are activated by multiple interactions between chromatin remodelling complexes, transcription factors, domain activation and nuclear location i.e. whether the gene is in the vicinity of heterochromatin or euchromatin. It is difficult to subdivide these levels of gene regulation as it is becoming increasingly apparent that these processes overlap, and even for gene repression it might be the absence of chromatin activators which ensures their silencing. In particular, the regulation of gene expression by domain activation is very poorly understood even though this is probably a critical determinant in the hierarchy of gene regulation. Some elements have been identified which may be involved in domain regulation but

new results suggests that the roles of these elements are still not understood (Grosveld, 1999).

1.5.1 Gene activation

The molecular basis for the decondensation of the higher order chromatin fibre *in vivo* is obscure. Early studies, separated chromatin into active and inactive regions by digestion with nucleases and the selective precipitation of active genes (Gottesfeld et al., 1975; Bloom and Anderson, 1978). Subsequent studies on the active and inactive fractions by microscopy, nuclease digestion and circular dichroism (Weintraub and Groudine, 1976; Bloom and Anderson, 1979) suggested that active chromatin had a more open conformation. We now know that active gene sequences are characterised by and are associated with a number of different features. These include hyperacetylation, hypomethylation, DNaseI sensitivity, linker histone depletion and localisation to transcriptionally permissive regions of the nucleus which would have given the impression, in early studies, of the chromatin fibre having a more open conformation. Detailed studies of the higher order chromatin fibre in our laboratory originally showed that there was a difference in the sedimentation rate of the active chicken erythrocyte β-globin gene compared to the inactive chicken ovalbumin gene in support of the early data (Kimura et al., 1983). However, further studies on this region (Fisher and Felsenfeld, 1986; Caplan et al., 1987) showed that the alteration in sedimentation rate was due to the presence of a hypersensitive site, and subsequent loss of a nucleosome, altering the mass and conformational flexibility of the fragment. Therefore in conjunction with other studies it has not been shown that there is a difference in higher order conformation between active and inactive genes, or between euchromatin and heterochromatin (Section 1.2.2).

Transcriptionally active genes are partially depleted of linker histones (Tazi and Bird, 1990; Kamakaka and Thomas, 1990) suggesting that linker histones act as general repressors of gene expression. This was consistent with the model for the formation of the higher order chromatin fibre where it is proposed that linker histones are responsible for the folding of the 30nm fibre such that depletion of linker histones might lead to chromatin unfolding (Thomas and Padayatty, 1984). However, this

view has been reconsidered as histone H1 can be depleted with very few detrimental effects, and is absent during the developmental stages of *Xenopus*. H1s are now thought to function as specific gene repressors (Panetta et al., 1998; Crane-Robinson, 1999) whereas specialised linker histones (Section 1.1.7) like H1° or H5 bind to chromatin independently of DNA sequence, and probably function as general repressors of transcription in active gene regions.

Active genes are associated with an absence of DNA methylation although the 5' domains of many genes, including all known housekeeping genes, are within CpG islands. These CpG islands are associated with a lower level of H1 compared to the bulk chromatin, as is transcriptionally active chromatin (Kamakaka and Thomas, 1990), and core histones H3 and H4 are heavily acetylated (Tazi and Bird, 1990). The role of CpG islands is unknown but they maybe responsible for marking active genes to ensure they remain in transcriptionally permissive regions of the nucleus. As mentioned previously, histone acetylation is associated with active genes. In the chicken β-globin locus both the β^{ρ} gene and the β^{A} gene are found to be highly enriched in acetylated nucleosomes fractionated from both 5 day and 15 day erythrocytes suggesting that a 'poised' gene is acetylated to a similar level as a transcriptionally active gene (Hebbes et al., 1992). In contrast, a transcriptionally inactive gene like ovalbumin is found to be under-acetylated (Hebbes et al., 1992).

1.5.2 Chromatin remodelling complexes

At a fundamental level nucleosomes create a topological problem for both transcriptional initiation and elongation, as a megadalton RNA polymerase must pass around the DNA which is associated with the histone octamer (Bednar et al., 1999). *In vitro* studies have shown that eukaryotic polymerases are either significantly slowed or stopped by a nucleosome. Transcription therefore requires nucleosome displacement, probably involving multiple steps, including nucleosome destabilisation followed by histone transfer onto adjacent molecules or gene regions (Adams and Workman, 1993). Two activities have been identified which serve to alter the underlying chromatin structure (Workman and Kingston, 1998) to facilitate transcription and other cellular processes such as V(D)J recombination (Cherry and

Baltimore, 1999). The first is multi-subunit nucleosome remodelling complexes such as SWI/SNF, RSC, NURF, and CHRAC which have been shown to increase the accessibility of the DNA, presumably by modifications of the nucleosomal structure which is required for the transcriptional activation of genes. The second type of complex has been shown to be involved in altering the acetylation state of nucleosomal histones (Struhl, 1998; Kingston and Narlikar, 1999).

The SWI/SNF complex (Peterson and Tamkun, 1995) is required for activating certain genes in yeast, and is thought to counteract a repressive chromatin structure. In yeast, this complex is thought to be targeted to gene-specific loci by transcriptional repressors such as Hir1p and Hir2p (Dimova et al., 1999). *In vitro*, SWI/SNF uses the energy of ATP hydrolysis to alter DNA-histone contacts to make nucleosomal DNA more accessible to transcription factors and restriction enzymes. Similar complexes have been identified in *Drosophila* and human cells. The exact mechanism whereby this family of proteins is able to modulate the chromatin conformation is unclear although it appears that histone tails are not required (Guyon et al., 1999). BAF57, a component of a SWI/SNF complex, contains an HMG domain (Zhao et al., 1998) and Brm/SNF2α is tethered to chromatin by an I/Y-like DNA binding domain (Bourachot et al., 1999) suggesting that these complexes contain components which allow them to associate with the DNA. A SWI/SNF related complex, CHRAC (Alexiadis et al., 1998), induces movements of intact histone octamers to neighbouring DNA segments, and is required for remodelling chromatin *in vitro* around origins of DNA replication (Alexiadis et al., 1998). As with other remodelling complexes RSC is capable of disrupting the nucleosome completely and transferring the histone octamer to another molecule of DNA (Lorch et al., 1999). The targetting mechanisms of these remodelling complexes are unknown although it is likely that they are recruited by specific transcription factors which in turn are targeted to specific DNA sequences in and around promoter regions.

Histone acetyltransferases (HATs) are found associated with active genes. The first HAT, Gcn5, was identified in yeast where it functions as part of a complex of

proteins connecting activators to the basal transcription machinery. *In vitro*, GCN5-dependent HAT complexes acetylate nucleosomal histones and can promote transcription from nucleosomal templates (Steger et al., 1998; Utley et al., 1998) (Ikeda et al., 1999). A number of proteins complexes are thought to alter the acetylation state of nucleosomes and it has now been shown that a subunit of the RNA polymerase II holoenzyme contains HAT activity (Wittschieben et al., 1999).

In contrast to nucleosome remodelling by SWI/SNF complexes which might be reset spontaneously, it is necessary for acetylated histones to be de-acetylated after gene activation in an active process. Consequently, a family of histone deacetylases (HDACs) have been identified that are required for nucleosomal condensation and transcriptional repression (Wolffe, 1996), both after transcriptional activation and for the stable repression of inactive gene sequences. Two classes have been identified in yeast: class I is represented by yeast Rpd3-like proteins and the second by Hda1 - like proteins. Currently, five yeast HDACs have been identified (Rundlett et al., 1996). In humans, both classes of HDAC have also been found: class I members are HDAC1, HDAC2, HDAC3 and class II members which were recently identified by using the amino acid sequence of yeast Hda1p to search the Genbank/EST database are HDAC4, HDAC5 and HDAC6 (Grozinger et al., 1999).

Although nucleosome remodelling and histone acetylation have been considered independently it is becoming clear that the large chromatin remodelling complexes act in concert with both SWI/SNF and HAT/HDAC activities. One of the yeast genes whose expression requires both GCN5 and SWI/SNF is HO (Perez-Martin and Johnson, 1998), which encodes a site-specific endonuclease that initiates the process of mating-type conversion (Pollard and Peterson, 1997). Expression of the yeast HO gene in late G1 of the cell cycle requires the SWI/SNF chromatin remodelling complex, the Gcn5p histone acetyltransferase, and two different sequence-specific transcriptional activators Swi5p and Swi4p/Swi6p. The Swi5 activator is required for recruitment of both the SWI/SNF nucleosome remodelling complex and the SAGA histone acetylase complex to the HO promoter (Cosma et al., 1999). Swi5-dependent histone acetylation of the HO promoter occurs prior to, and independently of

transcription (Krebs et al., 1999): for example, in the case of the yeast HO locus it appears there is a sequential order of activation in which the Swi5p activator recruits SWI/SNF whose remodelling activity then targets a Gcn5p-containing HAT complex to the promoter. Acetylation of the nucleosomes that encompass the cell cycle boxes may then facilitate binding of the Swi4p/Swi6p complex, which acts very late in the activation cycle. Finally, the Sin3p/Rpd3p deacetylase complex may play a significant role in erasing acetylation events, ensuring that the HO locus has a low level of histone acetylation prior to the beginning of the next cell cycle.

The SWI/SNF and GCN5 complexes are thought to function at specific genes to regulate their expression. In contrast, the NuRD (nucleosome remodelling and deacetylating) complex probably has a more global role in gene inactivation. This complex has both ATP-dependent nucleosome remodelling activity and deacetylase activity, presumably so that deacetylation can be facilitated by the remodelling activity. The role of this complex is still very unclear but as mentioned earlier (Section 1.2.6), it has been associated with Ikaros/Helios-dependent inactivation of gene sequences localised to regions of centromeric heterochromatin (Brown et al., 1999). Like Ikaros, BRCA1 encodes a zinc finger protein of unknown function and is able to interact via its carboxy terminus domain with HDAC1 and HDAC2 and possibly acts as an HDAC recruiter in an analogous fashion.

Heterochromatin is generally methylated and satellite sequences are bound by methyl-CpG-binding proteins MeCP1 (Cross et al., 1997) and MeCP2 (Meehan et al., 1989). MeCP2 has been shown to interact with HDAC complexes (Jones et al., 1998; Nan et al., 1998a). MBD2 (Hendrich and Bird, 1998), another methyl-CpG-binding proteins, is associated with NuRD (Zhang et al., 1999). These results suggest that CpG methylation, which has been associated with long-term gene silencing (Nan et al., 1998b), is mediated by the recruitment of a class of deacetylases, and which in turn silence specific genes by their recruitment to regions of repression by proteins such as Ikaros and possibly BRCA1.

1.5.3 Gene repression

Although the transcriptional ground state of eukaryotes is generally repressive, under certain situations it is necessary to maintain and extend this repression to avoid the aberrant expression of genes. One family of proteins first identified in *Drosophila*, based on their capacity to influence position effect variegation, share a common domain involved in chromatin modification and organisation known as the chromodomain which is conserved from yeast to man. The Polycomb group (PcG) of chromodomain transcriptional repressors regulate the expression pattern of the certain developmental genes in *Drosophila*. PcG genes are involved in repression of the homeotic genes, whereas trithorax genes are involved in activation of these same genes. PcG proteins are not DNA binding proteins but they have preferred sites of entry into chromatin (Pirrotta and Rastelli, 1994). Transcriptional activation by GAL4 is blocked in chromatin that is under PcG repression (McCall and Bender, 1996). Mutations in these genes can be suppressed by mutations in genes of the SWI/SNF family and a complex termed PRC1 (Polycomb repressor complex 1) that contains the products of the PcG genes was found to compete with SWI/SNF for remodelling of nucleosomal arrays in an acetylation-independent fashion (Shao et al., 1999). This suggests that PcG mediated repression maybe occurring by some mechanism involving nucleosome remodelling although a functional interaction has not been demonstrated. Two broad classes of mechanism have been proposed to explain repression by PcG complexes. An earlier hypothesis suggests that certain PcG proteins physically spread along the chromatin to produce a repressed state wherever they are present (Paro, 1990) (Figure 1.12A), whilst a second hypothesis proposes that PcG proteins are localised to discrete sites, from where they are able to repress access to the DNA over a wide region in competition with SWI/SNF family members (Figure 1.12B).

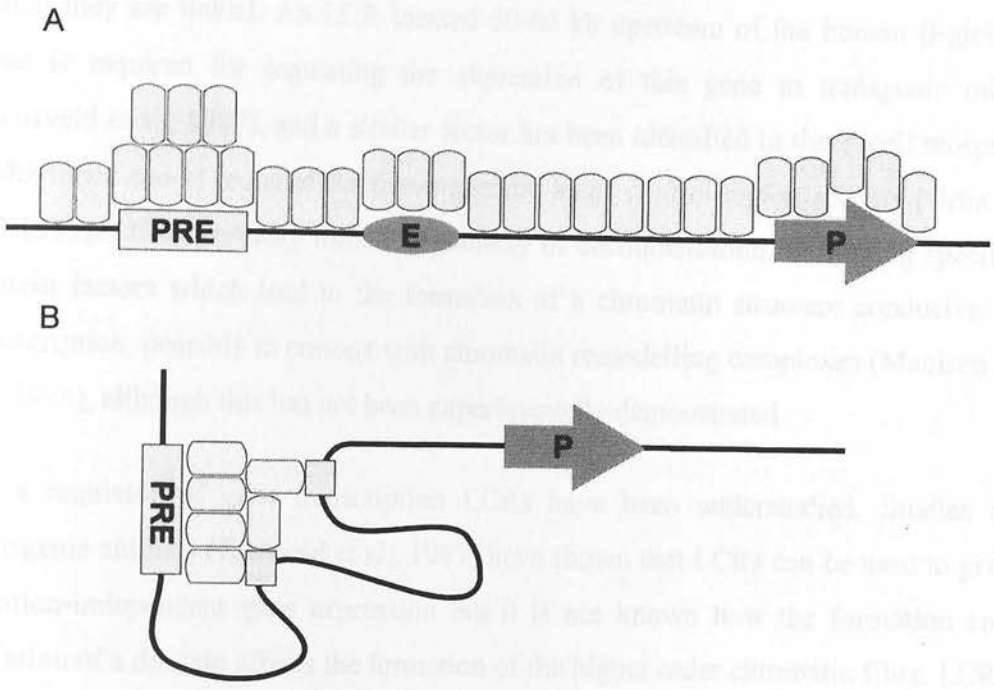


Figure 1.12 Polycomb group mediated gene silencing. (A) Spreading model. The PcG complex nucleates at the PRE and spreads by co-operative interactions with chromatin to involve enhancers (E) and promoters (P). (B) Hop and skip model. The PRE complex recruits and stabilises transient complexes formed at weak proto-PRE sites (yellow boxes) bringing the PcG complex into the vicinity of enhancers and promoters. (Redrawn from Pirrotta, 1998).

1.5.4 Boundary elements

Two types of DNA sequence which act in *cis* have been identified which may be involved in the unfolding of chromatin domains: locus control regions (LCR) and insulators (Bell and Felsenfeld, 1999).

LCRs (Kioussis and Festenstein, 1997; Higgs, 1998) are cis-acting regulatory elements thought to provide a tissue-specific open chromatin domain for genes to which they are linked. An LCR located 50-60 kb upstream of the human β-globin gene is required for regulating the expression of this gene in transgenic mice (Grosveld et al., 1987), and a similar factor has been identified in the T-cell receptor alpha locus and is required for tissue-specific locus control region activity (Ortiz et al., 1999). LCRs probably initiate a pathway of decondensation by binding specific protein factors which lead to the formation of a chromatin structure conducive to transcription, possibly in concert with chromatin remodelling complexes (Madisen et al., 1998), although this has not been experimentally demonstrated.

As a regulator of gene transcription LCRs have been understudied. Studies in transgenic animals (Grosveld et al., 1987) have shown that LCRs can be used to give position-independent gene expression but it is not known how the formation and isolation of a domain affects the formation of the higher order chromatin fibre. LCRs are frequently associated with DNaseI hypersensitive sites and are responsible for the temporal regulation of clustered tissue-specific genes by competing for transcription factors. Presumably these interactions will also have an influence on the formation of the higher order fibre, especially if the LCRs are looping to interact with other enhancer sequences.

Insulators are thought to be responsible for delimiting the activity of a domain (Udvardy, 1999). These elements are defined in an enhancer blocking assay where an enhancer is unable to interact with the adjacent promoter. The scs (specialised chromatin sequences) elements were identified at the *Drosophila* hsp70 locus and have been demonstrated to insulate genes from the activity of enhancers (Kellum and Schedl, 1992). The *Drosophila* proteins BEAF-32A and BEAF-32B bind to the scs'

boundary element of the hsp70 locus (Zhao et al., 1995; Hart et al., 1997; Cuvier et al., 1998). Another protein SBP (scs binding protein), has been shown to bind to a 24 base pair region of the scs element *in vitro* and *in vivo* (Gaszner et al., 1999). Although a role has not been associated with these novel proteins they appear to function in a co-operative fashion and maybe involved in the torsional regulation of the chromatin fibre in this region, especially as topoII is recruited to scs and scs' after heat shock induction. Recently an eleven zinc-finger protein called CTCF has been found to interact with a 42 bp fragment of the chicken β-globin insulator, and it has also been shown to bind to all other vertebrate insulators (Bell et al., 1999). Although the proteins which interact with CTCF have not been identified it would be surprising, based on other recent studies, if the protein did not in some way interact with a chromatin remodelling complex.

1.6 Thesis Aims

The chromatin fibre has been studied in detail for over twenty years. Much is now known about the conformation of the nucleosome and recent data shows that there is a large number of remodelling complexes present in cells which are responsible for re-organising the repressive environment of the chromatin to regulate and facilitate transcription.

The 30nm fibre is unequivocally considered to be the default structure of the higher order chromatin, although very little is known about its precise conformation both *in vitro* and more also *in vivo*. This is due to insufficient applicable structural techniques for the study of such a large, and probably irregular, nucleoprotein complex. One of the few techniques which is able to discriminate between conformations of the higher order fibre are hydrodynamic studies. These techniques have been used successfully in this lab (Allan et al., 1981; Kimura et al., 1983; Allan et al., 1986; Caplan et al., 1987) to analyse the conformation of different chromatin fibres. However, to date it has not been demonstrated that there is any alteration in the higher order conformation of either active or inactive genes. During cell differentiation a number of different loci need to be de-repressed and re-repressed to allow this cellular process to occur in a co-ordinated fashion. I have undertaken a

detailed study (Chapter 3) of the higher order chromatin fibre from embryonal stem cells and embryonal carcinoma cells before and after they are differentiated. I have used a combination of nuclease accessibility studies and hydrodynamic techniques to explore the higher order conformation of these different fibres to establish how the conformation of the chromatin fibre alters during cell differentiation.

Centromeres are a unique region of the chromosome (Section 1.3). As mentioned, they are generally associated with an underlying array of complex repeats, although in some situations it is possible to identify neocentromeres which are not associated with such structures. Currently there is much interest in the development of artificial chromosomes as therapeutic delivery vehicles but to date it has not been possible to identify a suitable sequence(s) that is appropriate for the generation of such gene transfer vectors. A deeper understanding of structural and functional components of centromeres may help in the generation of mammalian artificial chromosomes (Willard, 1998). Using a hydrodynamic approach I have analysed the chromatin conformation around the centromeres of both mouse and human, and I have developed a system which will help identify the key components responsible for the formation of centromeres (Chapter 4). As an off-shoot, this study has also provided an approach for the identification of chromatin regions which have an altered conformation.

The complement of linker histones is altered during both embryonic development and cell differentiation and linker histones have been shown to both globally and locally regulate the expression of genes. To help disseminate the role of globally acting linker histones and histone H5 in particular, and to establish whether an alteration in the complement of linker histone proteins can alter cell differentiation, I have expressed histone H5 in embryonic stem cells (Chapter 5). Embryonic stem cells have the capacity to differentiate in to all other cell types and therefore provide an ideal model system for identifying alterations in differentiation pathways.

2.1 Reagents, stock solutions and buffers

All reagents were purchased from Sigma or BDH unless stated otherwise.

Acrylamide - 40% stock solutions of 19:1 and 29:1 acrylamide/bisacrylamide were purchased from National Diagnostics.

Agarose gel loading buffer - 5× stock as 40% (w/v) sucrose, 5× TBE, 0.25% (w/v) bromophenol blue, 0.25% xylene cyanol if required.

Ammonium persulphate prepared as 10% and 25% (w/v) solutions in water and stored at 4°C for up to 3 days.

Antibiotics - A 1000× stock of ampicillin and kanamycin was prepared by dissolving ampicillin sodium salt or kanamycin monosulphate in distilled water at 100 mg/ml and 50 mg/ml, respectively. Small aliquots were stored at –20°C.

Chloroform:IAA comprised of chloroform and iso-amyl alcohol mixed at a ratio of 24:1.

Complete culture medium comprised Glasgow modified Eagle's medium (GMEM, Gibco-BRL) supplemented with 10% fetal calf serum, 0.1 mM β-mercaptoethanol and non-essential amino acids.

Coomassie blue stain comprised 45% methanol, 10% glacial acetic acid and 0.05% Coomassie blue R250. To facilitate dissolution the stain was initially dissolved in the methanol.

Coomassie blue destain comprised 10% glacial acetic acid and 10% methanol. Raw silk was used to absorb the Coomassie blue dye.

Denhardt's solution was prepared as a 100× stock consisting of 2% (w/v) BSA, 2% (w/v) Ficoll 400, 2% (w/v) polyvinyl pyrrolidone and normally used at 5×.

DEPC-water was prepared by diluting diethyl pyrocarbonate to 0.1% (v/v) in distilled water with agitation, followed by incubation overnight at room temperature and then autoclaved.

DIA/LIF was produced by transfecting Cos-7 cells with a murine DIA/LIF expression plasmid (Smith, 1991). Conditioned medium containing the DIA/LIF was stored in small aliquots at –20°C.

Di-butyryl cAMP (but_2cAMP) was dissolved at 100 mM in distilled water, filter sterilised and stored at –20°C.

DNA cellulose solid was mixed with cellulose CF11 at a ratio of 1:1 (w/w), hydrated in TEP_{2M} and stored at 4°C.

DNA markers - λ HindIII (NEB), ØX174 RF DNA HaeIII (HT Biotechnology), pBR322 MspI (HT Biotechnology), 1 kb ladder (Promega), and 100 bp ladder (Promega) were dissolved at 500 µg/ml in 1× agarose gel loading buffer. 500 ng-1 µg was normally loaded per lane. A 20 bp ladder (Biorad) was used for sizing small fragments on native acrylamide gels.

DNaseI was dissolved at 5 mg/ml (12,200 units/ml) in 50% glycerol, 150 mM NaCl.

DTT was prepared by dissolving solid dithiothreitol in distilled water at 1 M and stored in small aliquots at –20°C.

EDTA - Ethylene diamine-tetraacetic acid (disodium salt) was dissolved at 0.5 M in distilled water and adjusted to pH 8.0 with NaOH.

EGTA - Ethylene glycol-bis[β-aminoethyl ether]-tetraacetic acid was dissolved at 100 mM in distilled water and adjusted to pH 8.0 with NaOH.

Ethidium bromide stock solution was prepared by dissolving ethidium bromide solid to 50 mg/ml in distilled water, filtered and stored in a light-proof bottle at room temperature.

Ficoll gel loading buffer was prepared as for agarose gel loading buffer except ficoll at 17.5% w/v, for 5×, replaced the sucrose.

FRB buffer - 20× stock comprised 0.4 M MOPS, 100 mM sodium acetate, 20 mM EDTA. It was stored in the dark and discarded if it turned dark yellow.

Glycogen from oysters (USB) was used as a DNA carrier. It was dissolved at 20 mg/ml in TE, aliquoted and stored at –20°C.

IPTG was prepared by dissolving solid isopropyl β-D-thioglucoside at 1 M in distilled water and stored in small aliquots at –20°C.

Leishman's fix and stain was prepared by dissolving 1.5 g Leishman's in 1 litre of methanol using a heated magnetic stirrer. The solution was filtered through Whatman paper and stored at room temperature.

Luria-Bertani (LB) broth (Miller formulae) comprised 10% NaCl, 10% bacto-tryptone (Difco) and 5% yeast extract (Difco). It was sterilised by autoclaving and stored at room temperature.

Lysolecithin was dissolved in distilled water at 1 mg/ml and stored in small aliquots at –20°C.

3-Methoxybenzamide (Aldrich) was dissolved in dimethyl sulphoxide at 1.5 M and stored in 1 ml aliquots at 4°C.

Micrococcal nuclease stop buffer comprised 2% SDS, 200 µg/ml proteinase K, 5 mM EDTA.

PBS - Dulbecco's PBS (without Ca^{2+} and Mg^{2+}). 10× stock comprised 100 mM phosphate, 1.37 M NaCl, 27 mM KCl (equates to 80 g NaCl, 2 g KCl, 11.5 g disodium hydrogen phosphate (Na_2HPO_4), and 2 g potassium dihydrogen phosphate (KH_2PO_4) per litre). For tissue culture, the solution was sterilised by autoclaving.

Phenol was prepared as follows. 250 g phenol solid (Fluka) was dissolved in 127 ml 2 M Tris-HCl (pH 7.5) and once settled the aqueous phase (bottom) was removed. To the phenol phase 55 ml 2 M Tris-HCl (pH 8.0), 13.75 ml m-cresol, 550 µl β-mercaptoethanol and 275 mg 8-hydroxyquinoline was added. The solution was mixed well, aliquoted and stored at –20°C. Buffered phenol chloroform was prepared by mixing phenol, chloroform and iso-amyl alcohol at a ratio of 25:24:1.

Phosphate buffers. Two types of phosphate buffer were used: (i) 1M sodium phosphate (with respect to sodium ions) comprised 71 g disodium hydrogen phosphate (Na_2HPO_4) and 3.9 ml orthophosphoric acid (H_3PO_4); and (ii) 1 M sodium phosphate pH 7.6 (with respect to phosphate ions) was 13 ml 1 M sodium dihydrogen phosphate ($NaH_2PO_4 \cdot 2H_2O$) and 87 ml 1 M disodium hydrogen phosphate (Na_2HPO_4).

PMSF - 400× stock as 100 mM phenyl-methyl-sulphonyl-fluoride dissolved in isopropanol.

Protein gel markers were Dalton Mark VII-L (Sigma) - a mixture of proteins with the following approximate molecular weights: 66 kDa, 45 kDa, 36 kDa, 29 kDa, 24 kDa, 20 kDa, and 14.2 kDa. Markers were solubilised in 1× SDS sample buffer and were stored in small aliquots at –20°C.

Proteinase K was dissolved at 50 mg/ml in 50 mM Tris-HCl (pH 7.5), 2.5 mM $CaCl_2$, 50% glycerol and stored at –20°C.

Radioactive nucleotides were purchased from Amersham as follows: [α-^{32}P]dCTP 3000 Ci/mmol, [γ-^{32}P]ATP 3000 Ci/mmol and [α-^{35}S]dATP 600 Ci/mmol.

Retinoic acid was dissolved at 1 mg/ml in dimethyl sulphoxide and stored in a light-proof box in small aliquots at –20°C.

RNA loading buffer - 1.5× stock prepared by mixing 750 µl formamide, 90 µl 37% formaldehyde, 75 µl glycerol, 75 µl 20× FRB and bromophenol blue to approximately 0.05%.

RNaseA was dissolved in distilled water to 250 mg/ml and stored in small aliquots at –20°C.

Salmon sperm DNA was dissolved at 10 mg/ml in TE overnight. After hydration and sonication, it was stored in small aliquots at –20°C.

SDS -10% and 20% (w/v) stocks prepared in distilled water.

SDS sample buffer - 2× stock as 125 mM Tris-HCl (pH 6.75), 4% (w/v) SDS, 10% (v/v) β-mercaptoethanol, 20% (v/v) glycerol and bromophenol blue to approximately 0.1%.

Sephadex G-25/G-50 solid was hydrated in TE buffer supplemented with 0.05% sodium azide and stored at room temperature.

Sequencing stop mix comprises 95% formamide, 20 mM EDTA, bromophenol blue and xylene cyanol.

Sodium acetate 3 M pH 5.2

SSC - 20× stock as 3M NaCl, 0.3 M sodium citrate.

Taq Polymerase was prepared according to Pluthero (1993) and stored in aliquots at –70°C. Working aliquots were stored at –20°C.

TBE - 20× stock as 1 M Tris-borate, 20 mM EDTA (equates to 121 g Tris base, 62 g boric acid, 7.4 g $Na_2EDTA{\cdot}2H_2O$ per litre).

TE buffer - 10 mM Tris-HCl (pH 7.6), 0.1 mM EDTA.

TEP buffer - TE supplemented with 250 µM PMSF

TEP salt buffers - TE supplemented with 250 µM PMSF and the required amount of salt. TEP_5, 5 mM NaCl; TEP_{20}, 20 mM NaCl; TEP_{50}, 50 mM NaCl; TEP_{80}, 80 mM NaCl; TEP_{2M}, 2 M NaCl.

TPE - 10× stock comprised 900 mM Tris-phosphate, 20 mM EDTA (equates to 109 g Tris base, 15.5 ml 85% orthophosphoric acid, 7.4 g $Na_2EDTA{\cdot}2H_2O$ per litre (Sambrook et al., 1989))

Tris-glycine running buffer - 5× stock as 125 mM Tris, 1.25 M glycine and 0.5% SDS.

Trypsin/EDTA comprised 0.25 mg/ml trypsin (Difco), 1 mM EDTA, 1% chicken serum in PBS filter-sterilised through a 0.2 µm filter and stored in 20 ml aliquots at –20°C.

TST comprised 150 mM NaCl, 10 mM Tris-HCl (pH 8.0), 0.05% Tween 20. Normally prepared as a 10× stock without Tween.

X-gal (5-bromo-4-chloro-3-indolyl β-D-galactopyranoside) was dissolved at 50 mg/ml in dimethyl formamide and stored in small aliquots at –20°C in a light-proof box.

2.2 *Escherichia coli* culture and manipulation

2.2.1 Culture Media

For liquid culture, bacterial cells were grown in Luria-Bertani (LB) broth. Solid media was prepared by supplementing LB broth with 1.5% agar. For plasmid selection, ampicillin or kanamycin was added as required. For blue/white selection of transformants, LB plates were supplemented with appropriate antibiotics, 0.5 mM IPTG, 40 μg/ml X-gal and stored in the dark.

2.2.2 Bacterial strains

DH5α and DH10B (Gibco-BRL) were used for routine cloning. Both strains are *EndA*, and *RecA* making them suitable for preparing high quality plasmid DNA and for maintaining large plasmids in an intact configuration. BL21 (DE3), B834(DE3) (Novagen) and JM109 (DE3) (Promega) were used for protein expression studies. These strains carry an IPTG-inducible T7 RNA polymerase for high-level protein expression from a T7-based expression plasmid.

2.2.3 Bacterial glycerols

An aliquot of an overnight culture was added to an equal volume of 40% glycerol, 60% LB and stored at –70°C.

2.2.4 Preparation and transformation of competent cells

Plasmid DNA can be transferred into bacterial cells using two different approaches, chemical and electrical transformation. Chemical transformation is less efficient than electrical transformation although the technique is simpler, cheaper, and more reliable. Routinely, chemical transformation will give a transformation efficiency of 1×10^6 cfu/μg, whereas electrical transformation can give a transformation efficiency of 1×10^9 cfu/μg. It was found that for optimum transformation efficiency, the competent cells should be prepared in a cold-room at 4°C on ice.

Chemical transformation

Chemically competent cells were prepared using $CaCl_2$ (Hanahan, 1983). An overnight bacterial culture was diluted 1:100 into 100 ml of LB. The culture was grown with good aeration to an OD_{600} of 0.4. The cells were cooled on ice for 30 minutes, collected by gentle centrifugation (3000 rpm, 10 minutes, 4°C in a benchtop centrifuge) in a 50 ml tube, resuspended in 50 ml CG buffer (50 mM $CaCl_2$, 20% (v/v) glycerol) and incubated on ice for 60 minutes, respun and resuspend in 5ml CG buffer. The cells were divided into small aliquots for immediate use or frozen in a dry ice/ethanol bath and stored at –70°C.

For transformation, a small amount of plasmid DNA in less than 10 µl TE or water was added to 100 µl of competent cells and incubated on ice for 15 minutes. The cells were heat-shocked for 2 minutes in a water bath at 42°C, diluted into 0.5 ml LB and incubated at 37°C for 40 minutes. Aliquots of the cells were spread onto LB plates supplemented with appropriate antibiotics.

Electrical transformation

Electro-competent cells were prepared according to (Tung and Chow, 1995). An overnight bacterial culture was diluted 1:100 into 100 ml of LB grown with good aeration to an OD_{600} of 0.6 and cooled on ice for 30 minutes, collected by centrifugation (5000 rpm, 10 minutes, 4°C in a benchtop centrifuge) in a 50 ml tube, and washed twice with ice-cold sterile 10% (v/v) glycerol (prepared fresh) and finally resuspended in 0.2 ml ice-cold GYT (10% glycerol, 0.125% yeast extract, 0.25% bacto-tryptone). Aliquots were either used immediately or frozen in a dry ice/ethanol bath and stored at –70°C.

For transformation, 5 µl of plasmid DNA was added to 45 µl of competent cells and transferred to a sterile electroporation cuvette (Hybaid, 1mm gap) and chilled on ice for 5 minutes. The cuvette was pulsed in an electroporator (Hybaid) and the cells transferred into 0.5ml LB and incubated, with agitation to avoid clumping, at 37°C for 40 minutes. Aliquots were spread onto LB plates supplemented with appropriate antibiotics.

2.3 DNA purification

2.3.1 Phenol/chloroform extraction and ethanol precipitation

The DNA solution was adjusted to 200-400 µl with water. If substantial amounts of protein were present SDS was added to 0.1% and proteinase K to 100 µg/ml and incubated at 37°C for 30 minutes. An equal volume of buffered phenol chloroform was added and mixed by inversion or vortexing. Phases were separated by centrifugation (12,000 g, 5 minutes, room temperature in a microcentrifuge), the aqueous phase was removed to a fresh tube and re-extracted using buffered phenol chloroform. After recovering the aqueous phase, the sample was chloroform extracted with chloroform:IAA to remove residual phenol.

The DNA was precipitated by adding sodium acetate (pH 5.5) to 0.3 M and 2-2.5 volumes of ethanol and incubated at –20°C. The DNA was collected by centrifugation (12,000 g, 15 minutes, room temperature in a microcentrifuge) and the pellet washed with 1 ml 70% ethanol to remove residual salt. The DNA was re-centrifuged (12,000 g, 5 minutes, room temperature in a microcentrifuge), the liquid removed and the pellet dried under vacuum. The DNA was resuspended in an appropriate volume of water or TE.

2.3.2 Gel extraction

DNA fragments were resolved by agarose gel electrophoresis and the required fragment was excised using a clean razor blade. Two methods were used for purifying DNA from gel slices.

Electroelution

The gel slice containing the DNA was placed inside a dialysis bag containing a small amount of 0.5× TBE. The DNA was eluted from the agarose gel slice by electrophoresis in 0.5× TBE. The buffer containing the DNA was removed from the dialysis bag, purified by phenol/chloroform extraction and ethanol precipitation.

Phenol freeze-fracture

Small DNA fragments were purified from gel slices by phenol/freeze fracture (Bewsey et al., 1991). The gel slice was finely chopped and placed in a 1.5 ml centrifuge tube. 0.4ml of phenol was added and the tube vortexed vigorously then rapidly frozen in liquid nitrogen, thawed and vortexed. A further 0.4 ml phenol was added and the tube was vortexed vigorously then rapidly frozen in liquid nitrogen and thawed. 0.4 ml TE was added and the sample vortexed and centrifuged (12,000g, 15 minutes, room temperature in a microcentrifuge) to separate the phases. The aqueous phase was removed to a clean tube and the DNA was phenol/chloroform extracted and ethanol precipitated.

2.4 Preparation and manipulation of recombinant DNA

2.4.1 Preparation of plasmid DNA

Small quantities of plasmid DNA (<10 μg) were isolated from bacteria (typically 2 ml of overnight culture) using a commercial kit according to the manufacturer's instructions (QIA-Prep, Qiagen). Larger quantities of plasmid DNA (≥150 μg) were obtained from 50 ml of overnight culture using a commercial kit according to the manufacturer's instructions (Midi-Prep, Qiagen).

2.4.2 Restriction enzyme digestion

DNA was cleaved using restriction enzymes according to the manufacturer's instructions. The products of DNA restriction digests were analysed directly by agarose gel electrophoresis or purified by phenol/chloroform extraction followed by ethanol precipitation and resuspension in water or TE.

2.4.3 Dephosphorylation of DNA fragments

DNA fragments generated by restriction enzyme digestion are 5'-phosphorylated. The 5'-phosphates can be removed by alkaline phosphatase to facilitate further manipulation such as 5'-end labelling (Section 2.5.1) by polynucleotide kinase or to inhibit self- ligation (Section 2.4.6).

To DNA resuspended at approximately 50 μg/ml in TE (supplemented with 10 mM $MgCl_2$ for shrimp alkaline phosphatase) calf intestinal (CIAP; MBI-Fermentas) or shrimp alkaline phosphatase (SAP; Amersham or Roche) was added to 1 unit per pmol DNA ends and incubated at 37°C for 30-60 minutes (for blunt or 5'-recessed ends CIAP is incubated at 50°C). Calf intestinal alkaline phosphatase was inactivated by phenol/chloroform extraction and the DNA recovered by ethanol precipitation; shrimp alkaline phosphatase can be heat-inactivated at 65°C for 15 minutes.

In DNA dephosphorylation reactions shrimp alkaline phosphatase (SAP) was found to dephosphorylate DNA less efficiently than calf intestinal alkaline phosphatase (CIAP). This prompted a short comparison of the activities of two SAPs and a single CIAP source by comparing the rate at which they hydrolysed 4-nitrophenyl phosphate. In a 3ml assay mix of 100mM glycine-NaOH (pH 9.5) and 1 mM 4-nitrophenyl phosphate 1 μl of neat or 1 μl of a 1:10 dilution of the enzyme was added and the change in absorbance monitored at 400 nm. In some assays the reaction mixes were supplemented with 10 mM $MgCl_2$ and/or 1 mM $ZnSO_4$ as CIAP has a requirement for Zn^{2+} (although as it is permanently bound to the enzyme and the storage buffer contains Zn^{2+}, it should not be necessary to supplement it) and SAP has a requirement for Mg^{2+}. The reaction rates were calculated and are shown in Table 2.1.

Although CIAP seems to work better in normal reaction conditions it is clear from Table 2.1 that all of the enzyme preparations have a similar activity if supplemented with the correct ions.

Table 2.1 The rate of 4-nitrophenyl phosphate hydrolysis by alkaline phosphatase manufactured by three different suppliers in the presence of ion supplements.

Enzyme	Ion supplement		Number of units in assay as determined by manufacturer	Reaction rate (µmol/min)
	10 mM $MgCl_2$	1 mM $ZnSO_4$		
CIAP (MBI-Fermentas)	-	-	0.2	0.45
SAP (Roche)	-	-	0.1	0.04
SAP (Amersham)	-	-	0.1	0.05
CIAP (MBI-Fermentas)	+	-	0.2	0.39
SAP (Roche)	+	-	0.1	0.18
SAP (Amersham)	+	-	0.1	0.18
CIAP (MBI-Fermentas)	+	+	0.2	0.55
SAP (Roche)	+	+	0.1	0.18
SAP (Amersham)	+	+	0.1	0.13
SAP (Roche)	+	+	1	0.23
SAP (Amersham)	+	+	1	0.13

2.4.4 Converting overhanging to blunt ends

Many type II restriction enzymes cut DNA to leave a staggered end. These ends need to be blunted to allow two such fragments with incompatible ends to be ligated and to facilitate the ligation of blunt adapters onto the ends of these fragments. 3'-recessed ends can be filled by Klenow DNA polymerase, and single-stranded extensions can be removed by mung bean nuclease.

For 3'-end filling the DNA was dissolved at 50 µg/ml in 1× Klenow buffer (10 mM Tris-HCl (pH 7.5), 5 mM $MgCl_2$, 7.5 mM dithiothreitol) supplemented with 33 µM of all four dNTPs. 1 unit Klenow polymerase (NEB) was added per µg DNA and incubated for 15 minutes at room temperature. The reaction was stopped by adding EDTA to 10 mM and the Klenow inactivated by heating at 75°C for 10 minutes.

For removal of single-stranded extensions, the DNA was suspended at 100 µg/ml in 1× mung bean nuclease buffer (50 mM sodium acetate (pH 5.0), 30 mM NaCl, 1 mM $ZnSO_4$). 1 unit mung bean nuclease was added per µg DNA and incubated for 30

minutes at 30°C. The enzyme was inactivated by phenol/chloroform extraction and the DNA recovered by ethanol precipitation.

2.4.5 Adapter annealing

Single-stranded oligomers to be used as adapters (see Section 7.2 for details) were purchased from MWG-Biotech and were dissolved at 100 µM in water. Adapter oligos were annealed in high salt. In a total volume of 100 µl, the following were combined: 10 µM each adapter oligo, 1 M NaCl, 25 mM EDTA, 50 mM Tris-HCl (pH 7.6). The mix was heated to 90°C for 5 minutes and cooled slowly to 50°C. The adapters were then desalted on a Sephadex G-25 spin column against water. Annealed adapters analysed on a TBE 15% polyacrylamide gel are shown in Figure 2.1.

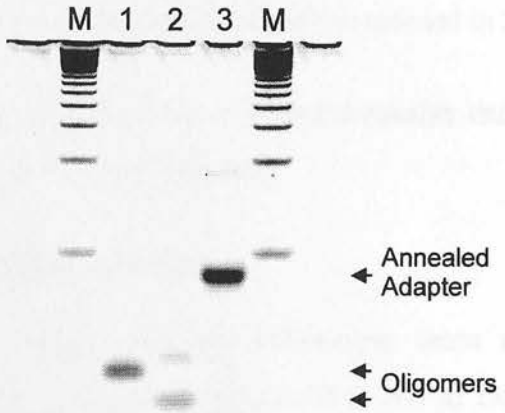


Figure 2.1 BstXI/Blunt adapter annealing. Adapters were analysed on a native 15% polyacrylamide gel to ensure the original oligomers were in good condition and the adapters had annealed efficiently. Lane 1 and 2 are the individual oligos (BstXI adapter 1 and BstXI adapter 2, respectively) and lane 3 is the annealed adapter. Markers are 20 bp ladder.

The oligomer pairs used in this project for preparing adapters are shown below:

Adapter Name		
EcoRI/BstXI	Adapter 1	5'-AATTCGGCCGCACTGGCCAGCACA-3'
	Adapter 2	5'-CTGGCCACTGCGGCCG-3'
BstXI/Blunt	Adapter 1	5'-GGCCGCACTGGCCAGCAC-3'
	Adapter 2	5'-CTGGCCAGTGCGGCC-3'

2.4.6 DNA ligation

DNA ligation reactions (20-50 µl) comprised of approximately 50 ng vector DNA added to insert DNA at a 1:3 molar ratio in 1× DNA ligase buffer (50 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, 25 µg/ml BSA) and 6 Weiss units of T4 DNA ligase. Sticky-end ligations were incubated at room temperature for 1 hr and blunt-end ligations at 16°C overnight. The ligated fragments were purified by phenol/chloroform extraction and ethanol precipitation in the presence of 20 µg glycogen as carrier. Samples were resuspended in 20 µl water.

Adapter to insert ligations were carried out in a similar manner except that a 1:100 molar ratio of DNA fragment to adapter was used.

2.4.7 Polymerase chain reaction

DNA fragments were amplified using the polymerase chain reaction (PCR). Routinely, preparative PCR reactions were performed in 100 µl 1× vent buffer (10 mM KCl, 20 mM Tris-HCl (pH 8.8), 10 mM $(NH_4)_2SO_4$, 2 mM $MgSO_4$, 0.1% Triton X-100) supplemented with 200 µM dNTPs, 0.5 µM primers (see Section 7.2 for details), 2 units Vent polymerase (NEB), and a small amount of template DNA. For preparative (as opposed to diagnostic) PCR reactions the number of cycles was reduced to 16 (from 25) to minimise the perpetuation of incorporation errors. Routinely, the PCR cycles comprised denaturation at 95°C for 1 minute, annealing temperature for 1 minute, at a temperature corresponding to the average melting temperature of the two primers, and extension at 74°C for 1 minute. The resulting PCR product was purified from an agarose gel by electroelution (Section 2.3.2).

2.4.8 DNA sequencing

The sequence of DNA was confirmed by the dideoxy chain-termination method (Sanger et al., 1977) either using fluorescent nucleotides with analysis on an ABI sequencer (Perkin-Elmer) or by incorporating [α-^{35}S]dATP with analysis by denaturing polyacrylamide gel electrophoresis and autoradiography.

Automated sequencing

Reactions were performed as recommended by Perkin Elmer using the ABI PRISM BigDye Terminator Cycle Sequencing Kit. Essentially, the DNA fragment was cycle-sequenced in the presence of fluorescent ddNTPs using a thermostable polymerase and heat denaturation.

Manual sequencing

Reactions were performed using the Sequenase kit as recommended by USB.

Template denaturation: 4-8 µg (approximately 1 pmol) double-stranded template DNA was suspended in 100 µl 0.2 M NaOH (e.g. 90 µl DNA and 10 µl 2 M NaOH) and incubated at room temperature for 10 minutes. The DNA was precipitated by adding 10 µl 3 M sodium acetate and 250 µl ethanol, incubated on ice for 10 minutes and then collected by centrifugation (12,000 g, 15 minutes, room temperature in a microcentrifuge). The pellet was washed in 70% ethanol and resuspended in 7 µl water.

Primer annealing: To the denatured template, 2 µl 5× Sequenase reaction buffer (200 mM Tris-HCl (pH 7.5), 100 mM $MgCl_2$, 250 mM NaCl) and 1 µl primer (approximately 1 pmol) were added. The primer and template were annealed by heating to 65°C for 2 minutes and cooled to 37°C over a 30 minute period.

Labelling reaction: The template was placed at 4°C in an ice/water bath and to each template, 5.5 µl elongation reaction mix (for 11 reactions use 11 µl labelling mix (7.5 µM dGTP, 7.5 µM dCTP, 7.5 µM dTTP) diluted 1 in 2.5, 11 µl DMSO, 11 µl 0.1 M

dithiothreitol, 5.5 µl [α-^{35}S]dATP, 19.5 µl Sequenase dilution buffer, and 2.5 µl Sequenase) was added, and incubated for 2-3 minutes.

Extension and termination reaction: To 2.5 µl each termination mix (8 µM ddNTP, 80 µM dGTP, 80 µM dATP, 80 µM dCTP, 80 µM dTTP, 50 mM NaCl) 3.5 µl labelling reaction was added and incubated at 39°C for 2-3 minutes. 4 µl sequencing stop mix was added to terminate the extension reaction. 1.5 µl of each sample was subsequently loaded onto a sequencing gel (Section 2.8.2).

2.5 Radioactive-labelling of DNA fragments

2.5.1 5'-end labelling

This technique was used for labelling oligomers and DNA size markers.

10 pmol of oligomer or 200 ng DNA size marker was suspended in 1× polynucleotide kinase buffer (50 mM Tris-HCl (pH 7.6), 10 mM $MgCl_2$, 10 mM β-mercaptoethanol) supplemented with 4 µl [γ-^{32}P]ATP and 10 units polynucleotide kinase in a total volume of 10 µl. The sample was incubated at 37°C for 45 minutes and the reaction stopped by heating to 68°C for 20 minutes. Unincorporated label was removed on a Sephadex G-25 or G-50 spin column (Section 2.5.3).

2.5.2 Random prime labelling

This technique was used for evenly labelling DNA fragments for use in Northern and Southern hybridisations (Feinberg and Vogelstein, 1983).

30ng DNA suspended in 15 µl water was boiled for 5 minutes and then chilled on ice. To this, 5 µl of labelling mix (50% (v/v) glycerol, 50 mM Tris-HCl (pH 7.6), 25 mM $MgCl_2$, 37.5 mM dithiothreitol, 0.125 mM dNTPs (except dCTP), 2 mg/ml BSA, 1.5 mg/ml random hexamers (Pharmacia)), 5 µl of [α-^{32}P]dCTP and 5 units Klenow polymerase were added. The sample was incubated at 37°C for 1 hr and unincorporated label was removed on a Sephadex G-50 spin column (Section 2.5.3).

2.5.3 Removal of unincorporated label

Unincorporated label was removed by passing the DNA sample through a sephadex spin column. Sephadex G-25 was used for oligomers and sephadex G-50 for other DNA fragments.

A 1ml syringe plugged with glass wool was filled with Sephadex slurry. Excess buffer was removed by centrifugation (2,000 rpm, 2 minutes in a swing-out benchtop centrifuge). If necessary, the column was washed through with water to remove residual buffer. The labelled sample was applied to the column in approximately 100 µl. The column was centrifuged (2,000 rpm, 2 minutes in a benchtop centrifuge) and the labelled DNA was recovered.

2.6 Cell culture

The cell lines used for this project were NIH3T3, Cos7, F9, 293 and three embryonic stem (ES) cell lines, ht2 (Section 3.1), MG1.19 (Section 5.3), and ZHTc6 (Section 5.3). Most of the techniques used apply to all of the cell types.

Methods for routine culture of cells were based on those described by (Smith, 1991), unless otherwise indicated. Cells were maintained in 5% CO_2 at 37°C in a humidified incubator. All tissue culture manipulations were undertaken in a laminar-flow sterile hood. To avoid bacterial or fungal contamination of the cell cultures, all objects and surfaces were sprayed with 70% industrial methylated spirits (IMS) before use.

2.6.1 Passaging cells

All media and solutions were prewarmed to 37°C in a water bath before use. Tissue culture grade flasks and dishes were gelatin-coated (0.1% in PBS), if required, for 20 minutes and aspirated before use. Culture medium was replenished every other day or when required and cells were subcultured when the flasks approached confluence.

The culture medium was aspirated and the cells were rinsed twice with PBS. Trypsin/EDTA solution was added to cover the cells and the cultures were incubated

at room temperature until the cells started to detach. The flasks were gently tapped to ensure complete dissociation of the cells from the surface and checked visually under a low-power magnification microscope. Complete culture medium was added to the flask to stop the trypsin digestion and the suspension was triturated to disaggregate cell clumps. The cells were transferred to a 20 ml universal and were pelleted by centrifugation (1200 rpm, 3 minutes, room temperature in a benchtop centrifuge). The cell pellet was resuspended in 5 ml fresh culture medium, counted on a hematocytometer if required, and were plated at a suitable density.

2.6.2 Freezing cells

Cells were harvested as described for routine passage. 5% DMSO was added to the cell suspension and the cells were aliquoted into freezing vials. The cells were initially frozen at –80°C for 24 hours before transferring to a liquid nitrogen cell store.

2.6.3 Thawing cells

Frozen vials were retrieved from liquid nitrogen storage and placed directly in a 37°C water bath. After thawing, the contents were transferred to a 20 ml universal containing 10ml medium. The cells were pelleted by centrifugation (1200 rpm, 5 minutes, room temperature in a benchtop centrifuge). The supernatant was aspirated and the cells resuspended in 10 ml complete culture medium and seeded in 25 cm^2 culture flasks.

2.6.4 Staining cells

Tissue culture cells were fixed and stained using Leishman's and ES cells were selectively stained for alkaline phosphatase activity. Cells were also stained for β-galactosidase activity to identify cells expressing the lacZ reporter gene.

Leishman's fix and stain

The tissue culture medium was aspirated and a small amount of stain was gently poured over the cells. After 2 minutes an equal volume of the stain was added and

after a further 2 minutes the stain was aspirated, the wells were washed with distilled water and allowed to air dry.

Alkaline phosphatase stain

The stain was purchased as a kit (Sigma, 86-R). Enough stain was prepared to cover each well evenly, the amounts below are suitable for a 6-well plate. Undifferentiated ES cell colonies stain intensely pink.

The stain was prepared fresh as follows. Mix 0.2 ml sodium nitrite in a small tube with 0.2 ml FRV-alkaline solution and allowed to stand for 2 minutes. This was added to 9 ml water in a 15 ml tube and mixed gently. To this 0.2 ml Napthol AS-BI alkaline solution was added.

The culture medium from the cells was replaced with 1.5 ml fix (for 100 ml: 25 ml citrate solution, 65 ml acetone, 8 ml formaldehyde) and left for 30 seconds. The fix was removed and the cells were rinsed gently with water and left for 45 seconds before rinsing further with water. 1.5 ml stain was gently poured into each well and the stain was incubated in the dark for 15 minutes at room temperature. The stain was removed and the wells were rinsed with water and air-dried.

X-gal staining

X-gal staining was used to stain cells which expressed the LacZ reporter gene.

Fix was 0.2% glutaraldehyde in PEM buffer (100 mM potassium phosphate (pH 7.4), 5 mM EDTA, 2 mM $MgCl_2$) and stain was 0.5 mg/ml X-gal, 10 mM potassium ferricyanide, 10 mM potassium ferrocyanide in PEM buffer.

The culture medium was aspirated and the cells were rinsed with PBS. The cells were fixed for 5 minutes and then washed 3×2 minutes in PEM buffer. Stain was poured over the cells and the cells were incubated overnight at 37°C until a blue colour could be seen. The cells were then stored at 4°C in PBS

2.6.5 *In vitro* differentiation of ES cells using 3-methoxybenzamide

ES cells can be induced to differentiate in monolayer culture by the withdrawal of DIA/LIF from the culture medium (Smith et al., 1988). However, at higher densities differentiation was heterogeneous and stem cell colonies can persist for a prolonged period. Specific and complete differentiation of ES cell monolayer cultures can be induced by exposure to 3-methoxybenzamide (Smith, 1991).

The specific effect of 3-methoxybenzamide (MBA) on ES cell differentiation is not affected by the presence of DIA/LIF in the culture medium. However, MBA used at levels marginally higher than needed for ES-cell differentiation causes extensive cell death. Therefore, the optimal concentration of MBA was titrated prior to induction (Figure 2.2). Leishman's fix and stain was used to stain all live cells, whilst wells in which the cells have died will not stain. Alkaline phosphatase stains stem cells, therefore wells containing stem cells will stain pink whilst differentiated cells will not stain. From the titration shown in Figure 2.2 it is clear that the optimal concentration of MBA was 6 mM. Since cell proliferation seems to be inhibited by MBA the differentiation of many cells was required to obtain sufficient material for further analysis.

For differentiation, the cells were passaged as described and 2×10^4 cells/cm^2 were seeded onto gelatine-coated flasks containing culture medium supplemented with 100 units/ml DIA/LIF. The ES cells were allowed to attach and grow overnight. The following day, MBA stock was diluted to 6mM in warmed culture medium. The ES cells were rinsed with PBS and the medium was changed to the MBA containing medium, supplemented with 10 units/ml DIA/LIF. The MBA containing medium was renewed after 24 hours to completely eliminate stem cells and then changed 24 hours later to culture medium without MBA or DIA/LIF. Differentiated cells were harvested when confluent (usually the following day).

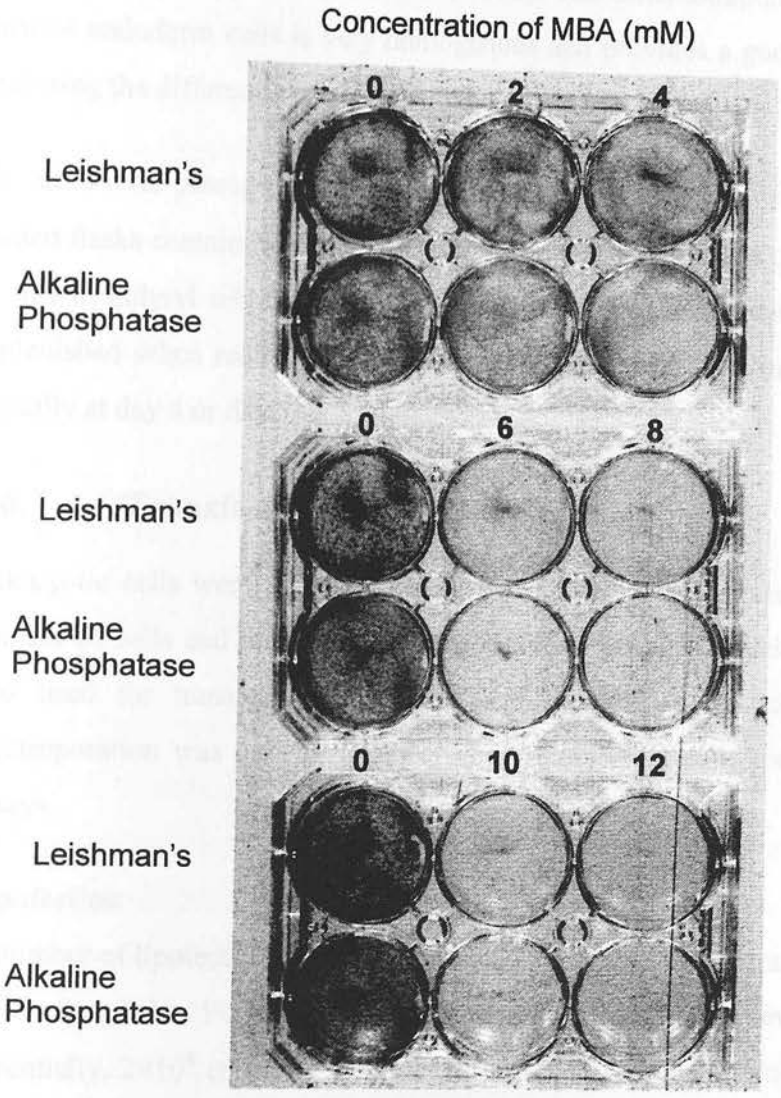


Figure 2.2 Differentiation of ES cells by 3-methoxybenzamide. ES cells were titrated with MBA to determine the optimal concentration of drug to induce differentiation. Leishman's stains all cells, whereas alkaline phosphatase stains stem cells. The optimal concentration of MBA was found to be 6 mM.

2.6.6 *In vitro* differentiation of F9 cells

F9 embryonal carcinoma cells have the capacity to differentiate into parietal endoderm-like cells by the addition of retinoic acid and di-butyryl cAMP (Strickland and Mahdavi, 1978; Strickland et al., 1980). The differentiation of F9 stem cells to parietal endoderm cells is very homogenous and provides a good model system for analysing the differentiation of cells in a defined manner.

The cells were passaged as described and 2×10^3 cells/cm^2 were seeded onto gelatine-coated flasks containing culture medium supplemented with 0.5 μM retinoic acid and 1 mM di-butyryl cAMP. The cells were cultured as normal and the medium was replenished when required. The differentiated cells were harvested when confluent (usually at day 4 or day 5).

2.6.7 Transfection of eukaryotic cells

Eukaryotic cells were transfected using two different techniques depending on the number of cells and the subsequent analysis that was to be carried out. Lipofection was used for transfecting small numbers of cells to test constructs, whereas electroporation was used to transfect large numbers of cells for colony formation assays.

Lipofection

A number of lipofection reagents were tested for their ability to transfect ES cells and it was found that FuGENE (Roche) consistently gave a high level of transfection. Essentially, 2×10^4 cells/cm^2 were plated in 6-well plates and grown overnight. The following day, a mixture of 2 μg linear or circular plasmid DNA and 5 μl Fugene was added to the culture medium and gently mixed.

NIH3T3, Cos7 and 293 cells were routinely transfected using lipofection at a density of 1×10^4 cells/cm^2. Cos7 and 293 cells were transfected in complete medium using 5 μl FuGENE (Roche) and 2 μg DNA per 10cm^2 dish. NIH3T3 cells were transfected in serum-free medium using 5 μl Lipofectamine (Gibco-BRL) and 1 μg DNA per 10

cm^2 dish. Three hours after transfecting NIH3T3 cells, the serum-free medium was replaced by complete medium.

Stable transfections

Clones were selected using a suitable concentration of selection drug (Table 2.2). The concentration of drug required to kill cells was determined empirically by plating cells in different drug concentrations and culturing for a suitable period of time until the cells at one concentration of the drug were all dead.

Table 2.2 Concentration of selection drug required to kill different cell lines.

Cell line	Drug	Concentration (μg/ml)
NIH3T3	G418	800
	Hygromycin B	200
Cos-7	G418	800
	Hygromycin B	200
ht2 (Section 3.x)	Ganciclovir	2.5 μM
	Hygromycin B	100
MG1.19 (Section 5.3)	G418	200
	Hygromycin B	100
ZHTc6	G418	200
	Tetracycline	1

To monitor transfection efficiency, control plasmids were frequently included in the transfection experiments. Early experiments used a β-galactosidase-expressing control plasmid (pPHCAGGS-lacZ, Plasmid #23 - See Section 7.1 for details) which could be monitored by X-gal staining (Section 2.6.4). Later experiments used a green fluorescent protein (GFP) expressing construct (pEGFP-C1 (Clontech), Plasmid #91 or pPHCAGGS-gfp, Plasmid #132), the expression of which could be monitored by fluorescence.

Electroporation

Electroporation of plasmids into ES cell was the routine method used in our laboratory for generating cells carrying a transgene. The use of an episomal expression system gives a higher level of transfection than normally found for electroporation and the expression constructs are rarely rearranged. The disadvantage of electroporation compared to lipofection is the large number of cells that are

needed. MG1.19 cells (Section 5.3) which support episomal replication of plasmids carrying a polyoma origin were transfected as follows: Cells plated the previous day were harvested as for routine passage and were washed twice in ice-cold PBS. The concentration of cells was adjusted to 6.25×10^6 cells/ml and 800 µl cells were placed into a pre-cooled electroporation cuvette (Biorad, 4 mm gap). 20 µg of plasmid DNA was added to the cells and mixed avoiding the formation of bubbles. The cells were electroporated (Biorad GenePulser) at 200 V, 960 µF giving a time-constant of approximately 20. After electroporation the cells were resuspended in 20 ml complete medium and plated at the following densities in 100 mm dishes - 5×10^4, 1×10^5, 2×10^5, 5×10^5 cells/cm^2. After 24 hours, the culture medium was changed and the cells were selected using appropriate drugs.

For the stable integration of targeted constructs into ES cells a slightly different method was used. The cells were resuspended at 2×10^7 cells/ml and 800µl of cells were transferred to a pre-cooled electroporation cuvette. 100 µg of linearised plasmid DNA was added to the cells and gently mixed. The cells were electroporated (800 V, 3 µF), resuspended in 10 ml complete medium and were plated at 1.5×10^6 cells per 100 mm dish. After 24 hours, the culture medium was changed and the cells were selected using appropriate drugs.

2.7 Preparations from tissue culture cells

2.7.1 Genomic DNA

Genomic DNA was typically prepared from 1×10^6-1×10^7 tissue culture cells. The cells were harvested as for routine passage and resuspended in 500 µl PBS. An equal volume of genomic lysis buffer (100 mM EDTA, 20 mM Tris-HCl (pH 7.6), 200 mM NaCl, 1% SDS, 200 µg/ml proteinase K) was added and the cells incubated overnight at 50°C. The DNA was purified by repeated phenol/chloroform extraction and the DNA was recovered by ethanol precipitation.

2.7.2 Total RNA

Total RNA was prepared using a modification of the technique of (Chomczynski and Sacchi, 1987). RNA was extracted using a phenol and guanidinium isothiocyanate solution (TRI Reagent, Sigma) as described by the manufacturer and precipitated using isopropanol. If required the RNA can be further purified to remove trace amounts of DNA by employing a selective LiCl precipitation step.

2.7.3 Nuclei

To prepare or analyse chromatin or nuclear proteins it was necessary to first prepare nuclei which can be subsequently processed as required. Cell nuclei were prepared using a modification of the method described by Cereghini and Yaniv (1984b).

Cells were harvested using trypsin/EDTA and washed in PBS/10% foetal calf serum. The cell pellet was resuspended in a small volume of NBA (85mM KCl, 10mM Tris-HCl (pH 7.6), 5.5% (w/v) sucrose, 0.5 mM spermidine, 0.2 mM EDTA, 0.25 mM PMSF). To this an equal volume of NBB (NBA plus 0.1% (v/v) NP40) was added and incubated on ice for 3 minutes, during which time the cell membrane lyses leaving the nucleus intact. The nuclei were collected by centrifugation (1200 rpm, 3.5 minutes, 4°C in a benchtop centrifuge) and washed in NBC (NBA minus EDTA) before resuspending at 20 A_{260} units/ml in NBC.

The nuclei concentration was determined by taking a small aliquot of the preparation, diluting 1:20 into NBC and adding a small amount of DNaseI. The sample was incubated at 37°C for 30 min and then further diluted 1:5 into sonication buffer (10 mM phosphate buffer, 5 M urea, 2 M NaCl). The absorbance was measured at 260 nm, and the concentration of the nuclei preparation was adjusted to 20 A_{260} units/ml.

2.7.4 Chromatin

Soluble chromatin was prepared by digesting freshly prepared nuclei with micrococcal nuclease (MNase) or a restriction enzyme and then lysing the nuclei overnight, releasing soluble chromatin.

Micrococcal nuclease

To prepare chromatin using MNase: Nuclei suspended in NBC buffer were supplemented with 1.5mM $CaCl_2$ and digested with 20-60 units MNase per 20 A_{260} units nuclei for 10 minutes at room temperature in the presence of 100 µg/ml RNaseA. The digestion was stopped by adding EDTA to 10 mM. After digestion the nuclei were resuspended in TEP_{20} supplemented with 300 µg/ml lysolecithin and were incubated at 4°C overnight. Nuclear debris was removed by centrifugation (12,000 g, 5 min, room temperature in a microcentrifuge) leaving the soluble chromatin in the supernatant.

Restriction enzyme

To prepare chromatin using a restriction enzyme: Nuclei were suspended in NBC supplemented with 2 mM $MgCl_2$ and 0.1 mM EGTA and digested using 100 units restriction enzyme per 1 A_{260} unit chromatin in the presence of 100 µg/ml RNaseA. The digestion was stopped after 8 minutes for AluI, or 64 minutes for MvaI by adding EDTA to 10 mM. After digestion the nuclei were resuspended in TEP_{20} supplemented with 300 µg/ml lysolecithin and incubated at 4°C overnight. Nuclear debris was removed by centrifugation (12,000 g, 5 min, room temperature in a microcentrifuge) leaving the soluble chromatin in the supernatant.

Chromatin purification

Soluble chromatin can be purified by centrifuging onto a 50% sucrose cushion in TEP_{80}. In an SW55 (Beckman) tube, 1.5 ml 50% sucrose in TEP_{80} was poured in the bottom of the tube and 10% sucrose in TEP_{80} carefully layered on top. The chromatin sample in a volume of up to 1 ml was layered on top of the sucrose. The sample was centrifuged at 48,000 rpm in a SW55 rotor for 90 minutes and 500 µl fractions

collected from the gradient by upward displacement. The peak fractions containing the chromatin were stored at 4°C for further analysis.

2.8 DNA analysis

2.8.1 Agarose gel electrophoresis

DNA preparations (other than oligomers) were size fractionated by horizontal gel electrophoresis. Different buffers and types of agarose were used depending on the size of the DNA fragments to be resolved.

(i) Fragments smaller than 100 bp were resuspended in 1× agarose gel loading buffer and resolved by electrophoresis in 2-3% NuSieve 3:1 (FMC) agarose gels in 1× TBE buffer.

(ii) Fragments larger than 100 bp and smaller than 1 kb were resuspended in 1× agarose gel loading buffer and resolved by electrophoresis in 1-1.5% standard agarose in 1× TBE buffer.

(iii) Fragments larger than 1 kb were resuspended in 1× ficoll loading buffer and resolved by electrophoresis in 0.7-1% standard agarose in 1× TPE buffer. Tris-phosphate was found to give better separation of DNA fragments larger than 3 kb compared to Tris-borate (Figure 2.3A) with good resolution over a wide range of DNA fragment sizes (Figure 2.3B). This buffer system was a modification of the E-buffer system described by Loenig (1969); it was found that for optimal results the buffer should be recirculated.

Gels were electrophoresed in 1× TBE or TPE supplemented with 0.5 μg/ml ethidium bromide at a constant voltage until the desired resolution was achieved. The DNA was visualised using a 320 nm transilluminator.

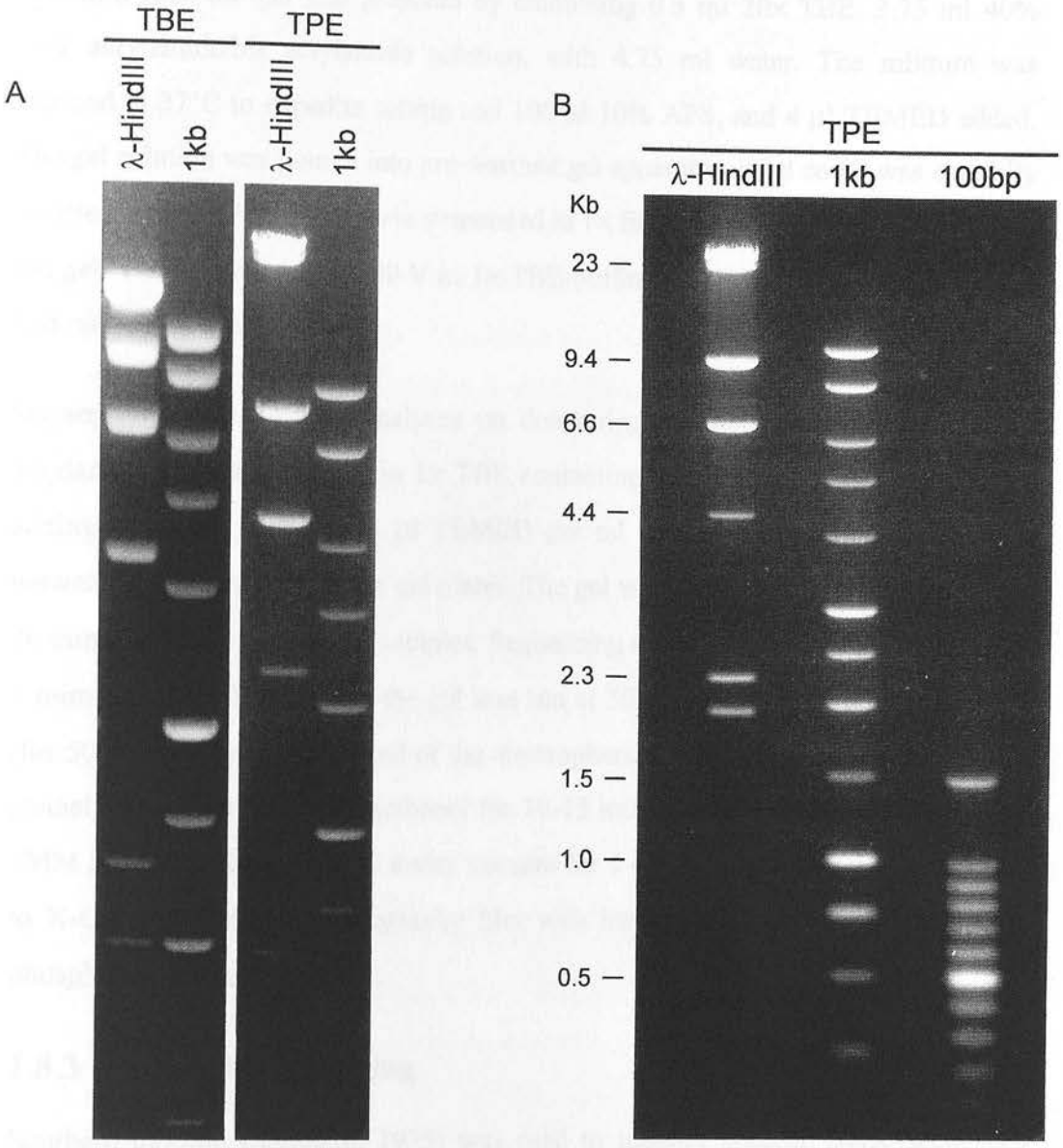


Figure 2.3 TPE gels resolve larger DNA fragments more clearly than TBE gels. (A) λ HindIII and 1 kb ladder were analysed on a 0.7% agarose gels at 1.5 V/cm for 20 hours under identical conditions except one gel used a TBE buffer system and the other a TPE buffer system. (B) λ HindIII, 1 kb and 100 bp ladders run on a 0.7% agarose gel at 2 V/cm for 12 hours.

2.8.2 Acrylamide gel electrophoresis

Oligomers were analysed on non-denaturing 15% polyacrylamide gels in 1× TBE. For mini-gels, the gel was prepared by combining 0.5 ml 20× TBE, 3.75 ml 40% 29:1 acrylamide/bis acrylamide solution, with 4.75 ml water. The mixture was warmed to 37°C to expedite setting and 100 µl 10% APS, and 4 µl TEMED added. The gel solution was poured into pre-warmed gel apparatus and a comb was carefully inserted. The DNA samples were suspended in 1× ficoll loading buffer and applied to the gel. The gel was run at 100 V in 1× TBE buffer until the bromophenol blue dye had run just over half-way.

Sequencing reactions were analysed on denaturing 6% polyacrylamide gels (19:1 acrylamide to bisacrylamide) in 1× TBE containing 8 M urea. The gel was cast by adding 1 µl 25% APS and 1 µl TEMED per ml acrylamide solution and poured between assembled clean glass gel plates. The gel was pre-run in 1× TBE for at least 20 minutes before loading the samples. Sequencing reactions were heated to 95°C for 5 minutes before loading and the gel was run at 50W (for 48 × 20cm gels) or 80W (for 50 × 38cm gels). At the end of the electrophoresis run the gel was fixed in 10% glacial acetic acid and 12% methanol for 10-15 minutes, transferred onto Whatman 3MM paper and dried at 80°C under vacuum for 1 hour. The dried gel was exposed to X-Omat (Kodak) autoradiography film with intensifying screens at –70°C or a phosphor-imager screen (Fuji).

2.8.3 Southern blotting

Southern blotting (Southern, 1975) was used to identify specific DNA sequences within a population of DNA fragments. Three steps were involved:

(i) DNA size fractionation on an agarose gel

(ii) DNA transfer and immobilisation on an inert membrane

After electrophoresis the gel was exposed to UV irradiation (3 minutes on a transilluminator). If the DNA fragments were larger than 5 kb the gel was

depurinated in 0.25 M HCl for 20 minutes to facilitate DNA transfer. The gel was washed in 1.5 M NaCl, 0.5 M NaOH for 40 minutes followed by washing in 1.5 M NaCl, 0.5 M Tris-HCl (pH 7.6) for 40 minutes. The DNA was transferred to Hybond N (Amersham) by upward capillary transfer in 20× SSC. After overnight transfer the filter was rinsed in 2× SSC and the DNA immobilised on the membrane by UV irradiation (3 minutes on a transilluminator).

(iii) Hybridisation

The membrane was placed in a hybridisation bottle with a mesh support and pre-hybridised with C&G buffer (0.5 M sodium phosphate (with respect to sodium ions), 7% (w/v) SDS, 1 mM EDTA (Church and Gilbert, 1984)) supplemented with 2% Marvel and 100 µg/ml salmon sperm DNA for 30 minutes at 65°C. 1-2 µg labelled probe was added per ml of pre-hybridisation buffer and hybridisation continued overnight at 65°C. The following day, the membrane was typically washed for 2 × 20 minutes with 0.5 M sodium phosphate, 1% SDS and 2 × 20 minutes 50 mM sodium phosphate, 1% SDS. Excess buffer was wiped from the membrane and the membrane exposed to X-Omat (Kodak) autoradiography film at –70°C with intensifying screens or a phosphorimager screen (Fuji).

2.8.4 Dot blotting

Dot blotting was used to determine the relative amounts of a specific DNA species in a number of different DNA samples. The DNA samples containing the DNA species of interest were spotted onto a nylon membrane using a vacuum manifold, denatured, immobilised and hybridised to a specific DNA probe. A critical step in dot blotting is to ensure that the DNA samples are fully denatured. Therefore, the approach used employs both alkali denaturation of the samples and alkali transfer to a positively-charged membrane.

The vacuum manifold was assembled using positively-charged Zeta-probe membrane (Biorad). The DNA samples were suspended in 200 µl 0.4 M NaOH/10 mM EDTA and heat denatured in a boiling water bath for 10 minutes. The wells in the vacuum manifold were flushed using water and the samples applied. The wells were rinsed

with 0.4 M NaOH and the vacuum manifold apparatus disassembled. The membrane was rinsed in 2× SSC, air-dried and the DNA was immobilised onto the membrane by UV irradiation (3 minutes on a transilluminator). The membrane was hybridised as for Southern blotting.

2.9 RNA analysis

2.9.1 Agarose gel electrophoresis

RNA was size fractionated in denaturing 1% agarose gels. RNA gels were prepared by adding agarose to 1× FRB buffer (Section 2.1) in DEPC-treated water and boiling. After the solution had cooled to 55°C, formaldehyde was added to 0.66 M and the gel was poured into a gel tank (pre-washed with 3% (v/v) hydrogen peroxide and rinsed with DEPC-water). RNA samples were prepared by adding 5 µl sample to 10 µl 1.5× RNA loading buffer. The samples were heated to 66°C for 15 minutes before loading and 1 µl of 10 mg/ml DEPC-EtBr was added. Samples were loaded and electrophoresis was performed at 80 V in 1× FRB buffer.

2.9.2 Northern blotting

Northern blotting was used to identify specific RNA sequences within a population of RNA fragments. Three steps were involved:

(i) RNA size fractionation on an agarose gel

(ii) RNA transfer and immobilisation on an inert membrane

Agarose gels were rinsed in DEPC-water and transferred onto Hybond N by upwards capillary transfer in 20× SSC. After overnight transfer, the membrane was rinsed in 2× SSC and the RNA immobilised onto the membrane by 3 minutes UV irradiation on a transilluminator.

(iii) Hybridisation (Krumlauf et al., 1987)

Hybridisation buffer was 60% formamide, 5× Denhardt's (Section 2.1), 5× SSC, 1% (w/v) SDS, 7% (w/v) dextran sulphate, 20 mM sodium phosphate (pH 6.8). The

buffer was warmed to facilitate dissolution, filtered through Whatman paper and salmon sperm DNA was added to 100 µg/ml.

The membrane was placed in a hybridisation bottle with a mesh support and pre-hybridised for at least 2 hours at 65°C in hybridisation buffer. 1-2 µg labelled probe was added per ml of pre-hybridisation buffer and hybridisation was continued overnight at 65°C. The following day, the membrane was washed for 2 × 20 minutes with 2 × SSC, 1% SDS and 2 × 20 minutes 0.1× SSC, 1% SDS at 65°C. Excess buffer was wiped from the membrane and the membrane exposed to X-Omat (Kodak) autoradiography film at –70°C with intensifying screens or a phosphor-imager screen (Fuji).

2.10 Protein Analysis

2.10.1 Protein Precipitation

Proteins were precipitated from solution using 6 volumes of acetone followed by centrifugation. The pellet was 90% acetone washed to remove residual salt and resuspended in TEP. Proteins were also precipitated from solution using 25% trichloroacetic acid.

A standard procedure for the selective isolation of linker histones is 5% perchloric acid (PCA) extraction, which characteristically partitions the linker histones into the soluble phase. The linker histones can be subsequently precipitated from the soluble phase using acetone or trichloroacetic acid.

2.10.2 Acrylamide gel electrophoresis

Proteins were routinely analysed by 15% SDS-PAGE in a Tris-glycine discontinuous buffer system (Sambrook et al., 1989).

Typically, 1 mm thick 9 × 9 cm slab gels were used. 10 ml of separating gel was made by mixing 5 ml 30% acrylamide (29:1, acrylamide: bisacrylamide), 2.5 ml 1.5 M Tris-HCl (pH 8.8), 2.4 ml water, 100 µl 10% SDS, 100 µl 10% APS and 4 µl TEMED. The gel solution was poured between the plates and a water overlay was

carefully applied. After setting (30 minutes) the water overlay was flushed and a 5% stacking gel was poured on top of the separating gel. 4ml stacking gel was prepared by mixing together 2.7 ml water, 0.67 ml 30% acrylamide, 0.5 ml 1.0 M Tris-HCl (pH 6.8), 40 µl 10% SDS, 40 µl 10% APS and 4 µl TEMED. After pouring the stacking gel a comb was placed in the top of the gel and allowed to set (5 minutes). The gel was mounted in the electrophoresis apparatus in Tris-glycine running buffer. The comb was removed and the wells were flushed.

Protein samples were suspended in 1× SDS loading buffer and boiled for 5 minutes before loading. Protein gels were run at a constant 25 mA for about 90 minutes. After electrophoresis, the gel was removed from the apparatus and stained in Coomassie blue stain for 2-3 hours. The stain was poured away and the gel was destained overnight in Coomassie destain with a hank of silk to absorb the dye. Stained protein gels were documented using 35 mm film or a digital camera, both with an orange filter to improve contrast.

2.10.3 Western blotting

For subsequent antibody probing, protein gels were transferred onto Hybond C (Amersham) using semi-dry blotting.

Anode buffer: 20% methanol, 300 mM Tris-HCl (pH 10.4)
Cathode buffer: 20% methanol, 25 mM Tris-HCl (pH 9.4), 40 mM EACA (ε-amino-n-caproic acid)
The protein gel apparatus was disassembled and the gel was rinsed in distilled water to remove excess SDS. The gel was placed on Hybond C soaked in anode buffer and sandwiched between 3 sheets each side of Whatman 17Chr cut to the size of the gel and soaked in either anode buffer or cathode buffer. The gel was blotted at 0.25 A for 40 minutes and then the filter was rinsed in TST, sealed in Saran wrap and stored at 4°C.

For antibody probing the membrane was blocked in TST plus 4% marvel for 1 hour at room temperature. Primary antibody was added at a suitable dilution in TST plus

4% marvel and the blot was further incubated for 1 hour. The primary antibody was washed off by 3 × 5 minutes washes in TST and the secondary antibody was added at a suitable dilution in TST plus 4% marvel. After 1 hour incubation the secondary antibody was removed and the blot was washed by 3 × 5 minutes washes in TST. The signal was normally visualised using chemiluminescence (ECL, Amersham). An equal volume of each ECL solution was added to the membrane and the membrane was incubated for 1 minute at room temperature. The filter was exposed to BioMax-MR film (Kodak) and developed.

2.11 Chromatin analysis

2.11.1 Sucrose gradient fractionation

Soluble chromatin (Section 2.7.4) was fractionated using sucrose gradient sedimentation in TEP_{80}. 6-40% isokinetic sucrose gradients were prepared in SW41 tubes leaving 0.5 cm at the top of the gradient to load the sample. 400 µl chromatin was loaded per gradient and centrifuged (3 hrs, 48,000 rpm, 4°C) in an SW41 rotor. 500 µl fractions were collected from the gradients by upward displacement and were purified and analysed by gel electrophoresis or stored at 4°C for further analysis.

2.11.2 Micrococcal nuclease digestion of nuclei

Nuclei (Section 2.7.3) were diluted to 4 A_{260} units/ml in NBC supplemented with 1.5 mM $CaCl_2$. To a 1 ml aliquot, RNase A was added to 100 µg/ml and a 100 µl aliquot removed into MNase stop buffer for time 0. To the remaining 900 µl, of sample 45 units MNase were added and 100 µl aliquots were subsequently removed at suitable time points into MNase stop buffer. After digestion the samples were incubated at 37°C for 30 minutes to digest remaining protein and the samples were phenol/chloroform extracted and the DNA was recovered by ethanol precipitated. The samples were resuspended in 40 µl 1× ficoll loading buffer and 20 µl aliquots were analysed on a 1% agarose gel in TBE.

2.11.3 Restriction enzyme digestion of nuclei

Nuclei were diluted to 4 A_{260} units/ml in NBC supplemented with 2 mM $MgCl_2$ and 0.1 mM EGTA. To a 1ml aliquot, RNase A was added to 100 µg/ml and a 100 µl aliquot was removed into MNase stop buffer for time 0. To the remaining material 2 units/µg restriction enzyme was added and 100 µl aliquots were removed at suitable time points into MNase stop buffer. After digestion the samples were incubated at 37°C for 30 minutes to digest remaining protein and the samples were phenol/chloroform extracted and the DNA was recovered by ethanol precipitation. The samples were resuspended in 40 µl 1× ficoll loading buffer and 20 µl aliquots were analysed on a 1% agarose gel in TBE.

2.11.4 Micrococcal nuclease digestion of naked chromatin

A single fraction of soluble chromatin purified by sedimentation in a sucrose gradient was supplemented with 3 mM $CaCl_2$. To a 500 µl aliquot, RNase A was added to 100 µg/ml and a 80 µl aliquot was removed into MNase stop buffer for time 0. To the remaining material, 0.2 units/µg micrococcal nuclease was added and 80 µl aliquots were removed at suitable time points into MNase stop buffer. After MNase digestion the samples were incubated at 37°C for 30 minutes to digest remaining protein and the samples were phenol/chloroform extracted and the DNA was recovered by ethanol precipitation. The samples were resuspended in 1× ficoll loading buffer and were analysed on a 1% agarose gel in TBE.

2.11.5 DNaseI digest on naked chromatin

A single fraction of soluble chromatin purified by sedimentation in a sucrose gradient was supplemented with 3 mM $MgCl_2$ and 1.5 mM $CaCl_2$. To a 500 µl aliquot RNase A was added to 100 µg/ml and a 80 µl aliquot was removed into MNase stop buffer for time 0. To the remaining material, 0.2 units/µg DNaseI was added and 100 µl aliquots were removed at suitable time points into MNase stop buffer. After digestion the samples were incubated at 37°C for 30 minutes to digest remaining protein and the samples were phenol/chloroform extracted and the DNA was recovered by ethanol precipitation. The samples were resuspended in 1× ficoll loading buffer and were analysed on a 1% agarose gel in TBE.

3. Higher order chromatin structure of ES and F9 cells

3.1 Introduction

The conformation and nuclear organisation of the chromatin fibre is critical for the regulation of gene expression in higher eukaryotes (Section 1.5). This acute regulation is particularly important during embryonic development, where it is necessary to precisely regulate key cellular processes such as proliferation, lineage determination and differentiation. In the early embryo there are many populations of stem cells which have the capacity to differentiate into a variety of distinct daughter cells. Two features which are unique about stem cells is their capacity to self-renew (Ogawa, 1993) and upon induction by appropriate molecular signals (frequently cytokines, for example Heyworth et al., 1990), to initiate a program of cellular differentiation. The molecular mechanisms which underlie and regulate cellular differentiation have been extensively studied, but to date the contribution and conformation of the higher order chromatin architecture during this process has not been examined.

Previous studies have been unable to demonstrate an alteration in the conformation of the higher order chromatin fibre between active and inactive genes (Fisher and Felsenfeld, 1986; Caplan et al., 1987). So, in an attempt to determine whether changes in the organisation of the higher order chromatin fibre occur during and contribute to a key cellular process, I have studied the conformation of the higher order chromatin fibre in stem cells prior to and after differentiation. The cell types used for these experiments were F9 embryonal carcinoma cells differentiated into parietal endoderm-like cells and mouse ht2 embryonic stem cells which were differentiated into epidermal-like cells. In addition, I have used mouse NIH3T3 embryonic fibroblasts as a terminally differentiated cell type.

F9 cells are derived from a teratocarcinoma cell line (Bernstine et al., 1973) and show very little spontaneous differentiation *in vivo* or *in vitro* (Sherman and Miller, 1978). Upon induction with retinoic acid, these cells differentiate to form cells analogous to the primary endoderm of the 4.5-day mouse embryo (Strickland and

Mahdavi, 1978). Subsequently, these precursor cells can be differentiated into two distinct populations of extra-embryonic endoderm which normally appear in the mouse embryo shortly after implantation. Parietal endoderm cells are formed by further treatment of the precursor cells with di-butyryl cAMP (Solter et al., 1979; Strickland et al., 1980) whilst visceral endoderm is formed when these cells are grown at high density in embryoid bodies (Hogan et al., 1981). The differentiation of F9 cells is irreversible, mostly synchronous and fairly rapid (Strickland and Mahdavi, 1978) and differentiation of the cells results in the loss of their tumourigenicity (Martin, 1980). The extent to which these cells represent the *in vivo* situation is debatable as the cells have been taken out of their normal developmental context and are lacking essential modulating factors. However, for a model system, the accessibility and purity of the differentiated cells make the F9 cell line ideally suited to differentiation studies (Sleigh, 1992).

Retinoic acid exerts its effects, in part, by acting through two types of nuclear receptor, the retinoic acid receptors (RAR) and the retinoid X receptors, both of which are members of the nuclear receptor superfamily. Addition of retinoic acid to F9 cells that have had the RARβ isoform deleted do not exhibit a fully differentiated phenotype (Faria et al., 1999), suggesting that RARβ is a primary target in the differentiation process.

Mouse embryonic stem (ES) cells are pluripotential cells that have the capacity to differentiate into all other cell types including the germline (Bradley et al., 1984; Nagy et al., 1993). Primarily, these cells have been used for the formation of transgenic animals and for the generation of specific types of progenitor cells (Keller, 1995; Nakano et al., 1996; Dani et al., 1997; Li et al., 1998; Brustle et al., 1999; Cho et al., 1999). Importantly, pluripotential embryonic stem cells provide a good system for studying cellular differentiation pathways. These cells can be maintained in the stem cell state indefinitely in the presence of the cytokine LIF, through a mechanism of self-renewal. These cells have been shown to repopulate both early mouse and rat blastocysts and if transplanted into suitable pseudo-pregnant recipients can develop to term, producing chimeric offspring. It is relatively easy to generate ES cells from

certain mouse strains (Nichols et al., 1990), and the frequency of obtaining ES cells from other strains can be increased by culturing the epiblast in the absence of other blastocyst-derived tissues (Brook and Gardner, 1997). Surprisingly, it has not been possible to generate rat ES cells which suggests that there is some fundamental difference between early rat and mouse stem cells. Recently, human embryonic stem cells have been isolated although they are phenotypically quite different to mouse ES cells (Thomson et al., 1998; Shamblott et al., 1998). Mouse ES cells can be readily differentiated by withdrawing LIF from the culture medium (Smith et al., 1988 ; Section 2.6.5). However, when dealing with large numbers of cells, I have found that this approach gives a rather heterogeneous population of cells, so I have differentiated ES cells using 3-methoxy benzamide (MBA) which gives rapid homogenous differentiation (Smith, 1991; Section 2.6.5). The cell type of the daughter cells is not clearly defined although morphologically the cells are epidermal-like.

The POU domain transcription factor Oct-4 has been shown to be expressed in blastomeres, pluripotent early embryo cells and the germ cell lineage (Rosner et al., 1990; Scholer et al., 1990). *In vitro*, Oct-4 is only found in undifferentiated stem cells. Oct-4 has been identified as a critical determinant of the embryonic stem cell phenotype (Nichols et al., 1998) and is found to be down-regulated during the differentiation of stem cells. The ES cell line used in these experiments, ht2 (Hitoshi Niwa, Unpublished), is a CGR8 derivative which has been targeted at the Oct-4 locus with a gene that expresses a hygromycin-thymidine kinase fusion protein (Lupton et al., 1991). These cells are ideally suited to differentiation studies because stem cells which express Oct-4 can be selected using hygromycin, and differentiated cells which do not express Oct-4 can be selected for using ganciclovir, which will kill the remaining stem cells still expressing Oct-4. Therefore, using this cell line it is possible to obtain homogenous populations of both stem cells and differentiated cells.

It was originally thought that after a cell had differentiated along a certain pathway and had permanently switched certain gene-specific loci off, it was impossible to

reverse this process and reactivate these genes due to specific methylation, deacetylation or ordered nucleosome positioning patterns. However, in nuclear transfer experiments, it is possible to take a mature nucleus from an adult cell and transplant it into a suitably primed egg, and after re-introduction into a suitable recipient, generate an embryo (Wilmut et al., 1997; Wakayama et al., 1998). From this it is clear that although chromatin is responsible for a general repression of gene loci it is possible for cells to reactivate lineages to a more immature state. Few studies in the past have carefully analysed the chromatin structure found in stem cells and in differentiated cells derived from them, although current ideas suggest that the chromatin structure which is responsible for gene repression must also be dynamic, such that under particular circumstances a once committed cell can regress to a cell which has a greater capacity to differentiate.

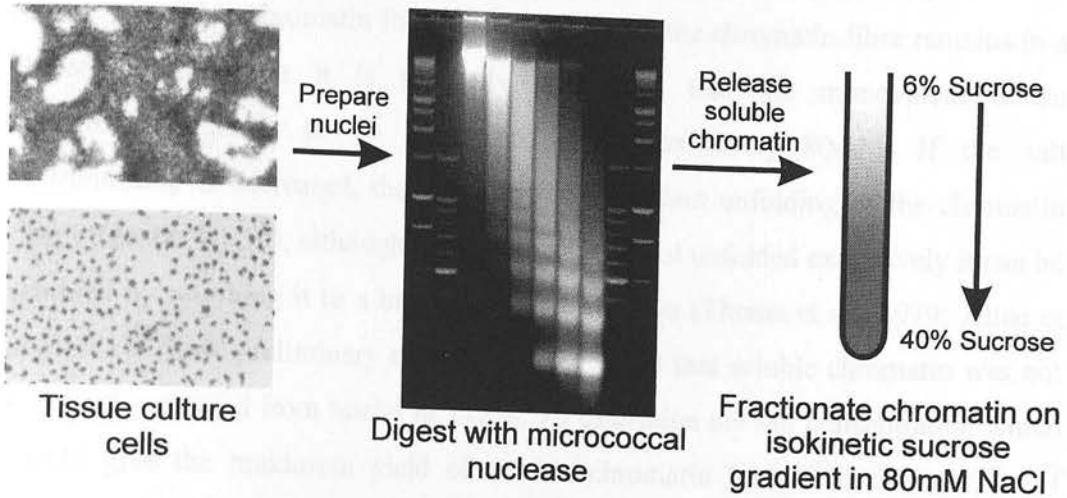
3.2 Techniques used for analysing the conformation of the higher order chromatin fibre

The biophysical techniques which have been used to study the higher order chromatin fibre include electron microscopy, fibre diffraction, neutron diffraction, electric dichroism and light scattering. Together these techniques have generated a substantial volume of results giving a partial picture of the chromatin fibre (Van Holde, 1988). Small angle neutron diffraction has been used to study the higher order chromatin fibre in nuclei (Baudy and Bram, 1979; Langmore and Schutt, 1980) and using this technique it has been possible to find differences in the conformation of the chromatin in nuclei with a high genetic activity, presumably because the fibres are less ordered (Notbohm, 1986). Recently, cryoelectron microscopy has been used to study chromatin fibres in cells, although this technique is unsuitable for studying the conformation of the bulk higher order chromatin fibre (Horowitz et al., 1997). Some studies have used nucleases to probe chromatin structure (for example, Huebner et al., 1981) but these approaches can lead to artifactual results due to the target of the nuclease being poorly defined.

However, the above techniques tend to be unsuitable or inaccessible for comparing the conformation of the higher order chromatin fibre from different cell types. One of

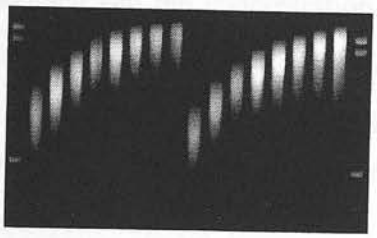
the few methods which is ideally suited to the study of different chromatin fibres is a hydrodynamic approach using sucrose sedimentation analysis of the samples. The basis for this technique is, if two chromatin fragments of the same mass have the same conformation they will sediment at the same rate in a sucrose gradient, but if the two chromatin fragments have different sizes or the same size but different conformations they will sediment at different rates. Unfortunately, due to the number of factors which could be involved in the structure of the chromatin fibre, it is not really feasible to identify what the specific changes in structure are between the two fibres in anything except in the most general terms. For convenience and versatility, I have chosen to use a hydrodynamic approach to compare chromatin fibres from different cell types (Allan et al., 1981; Allan et al., 1982a; Kimura et al., 1983; Allan et al., 1986; Caplan et al., 1987) as described below. In addition, I have used nucleases to probe the conformation of different chromatin fractions taken from sucrose gradients.

To prepare chromatin for analysis (see Section 2.7 for details), cell nuclei are briefly digested with either micrococcal nuclease or a restriction enzyme to generate large polynucleosome arrays (5-15 kb). The nuclei are lysed overnight to release the soluble chromatin, and the chromatin is analysed by isokinetic sucrose sedimentation analysis (Section 2.11.1). The chromatin fractions are pumped off the gradient by upward displacement and the DNA is purified (Section 2.3.1). After agarose gel electrophoresis (Section 2.8.1), a profile can be made which plots DNA fragment size versus the gradient fraction number, which is directly related to the extent of sample sedimentation in the gradient. This procedure is summarised in Figure 3.1. The shape and gradient of the plots gives an indication of the conformation of the chromatin fibre as two chromatin fibres of similar masses will only sediment at different rates if their conformations are different.

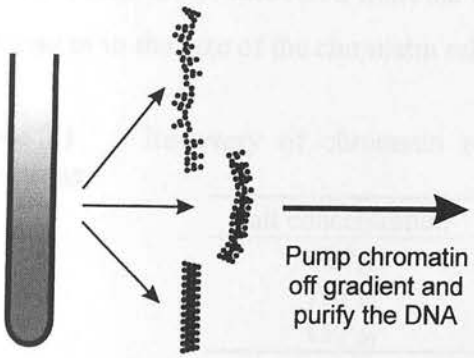


The sucrose gradient will fractionate chromatin based on

(i) Mass of the chromatin fragment
(ii) Conformation of the chromatin fragment



Fractionate DNA on agarose gel and plot gradient fraction number versus DNA size



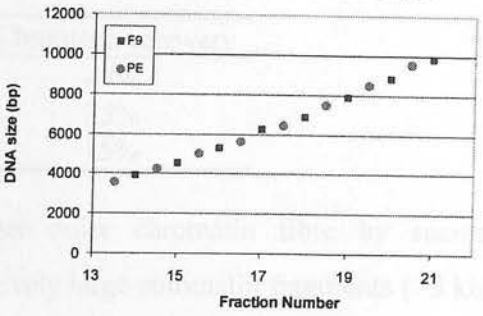


Figure 3.1 A hydrodynamic approach for analysing the conformation of the higher order chromatin fibre. Chromatin fragments are generated by digesting nuclei with micrococcal nuclease. Soluble polynucleosomes are released overnight and fractionated by size and conformation in an isokinetic sucrose gradient. Purification and analysis of the DNA taken from the gradient enables a relationship to be plotted of DNA size against the relative rate of chromatin sedimentation.

Preliminary experiments were used to alter a number of variables to optimise the release of soluble chromatin from nuclei. To ensure the chromatin fibre remains in a folded conformation it is generally considered that the monovalent cation concentration needs to be maintained at approximately 80mM. If the salt concentration is decreased, there is a cation-dependent unfolding of the chromatin fibre (Section 1.1.10), although if the chromatin is not unfolded extensively it can be refolded by returning it to a higher salt concentration (Thoma et al., 1979; Allan et al., 1981). These preliminary experiments suggested that soluble chromatin was not efficiently released from nuclei in TEP_{80}. To determine the salt concentration which would give the maximum yield of soluble chromatin I checked the amount of chromatin released from nuclei at different salt concentrations. Nuclei were prepared from F9 cells and the chromatin was digested with micrococcal nuclease. After digestion the nuclei were split into three aliquots and were gently pelleted and resuspended in either TEP_{80}, TEP_{50} or TEP_{20}. The samples were incubated on ice overnight and the following day the nuclei were pelleted and the concentration of soluble chromatin was measured (Section 2.7.4). From the results shown in Table 3.1 chromatin is released more efficiently in TEP_{20}, compared to TEP_{80}. From gel analysis of the DNA recovered from the soluble chromatins there were no apparent differences in the size of the chromatin released (data not shown).

Table 3.1 Recovery of chromatin released from nuclei under different salt conditions.

Salt concentration	Chromatin recovery
TEP_{20}	48%
TEP_{50}	25%
TEP_{80}	15%

To study the conformation of the higher order chromatin fibre by sucrose sedimentation it is necessary to analyse relatively large chromatin fragments (>5 kb). However, the release of large chromatin fragments from nuclei is less efficient than the release of small chromatin fragments, presumably because the fragments are more likely to aggregate. To examine the relationship between chromatin release and the size of the peak chromatin fragments, chromatin from F9 cells was prepared by digesting nuclei with different amounts of micrococcal nuclease. After digestion the

nuclei were pelleted, resuspended in TEP_{20} and the chromatin was released overnight. The following day the nuclei were pelleted and the concentration of the soluble chromatin was measured. The chromatin was loaded onto an isokinetic 6-40% sucrose gradient in TEP_{80} and the samples were spun for 4 hours at 41,000 rpm in an SW41 rotor (Section 2.11.1). The fractions were pumped off the gradients through a UV monitor at 254nm (Figure 3.2). The profiles show that with more extensive chromatin digestion, (Compare C to A in Figure 3.2) smaller chromatin fragments are released from nuclei more efficiently.

To release sufficient chromatin from nuclei to be able to undertake subsequent experiments but to ensure the chromatin fragments are large enough for the analysis it is necessary to compromise on the conditions used. For subsequent experiments, I have released chromatin from nuclei in TEP_{20} and allowed it to refold in the sucrose gradients in TEP_{80} buffer. To analyse suitably sized chromatin fragments chromatin was digested using 30-40 units micrococcal nuclease per 20 A_{260} units/ml of nuclei (equivalent to Figure 3.2 between A and B).

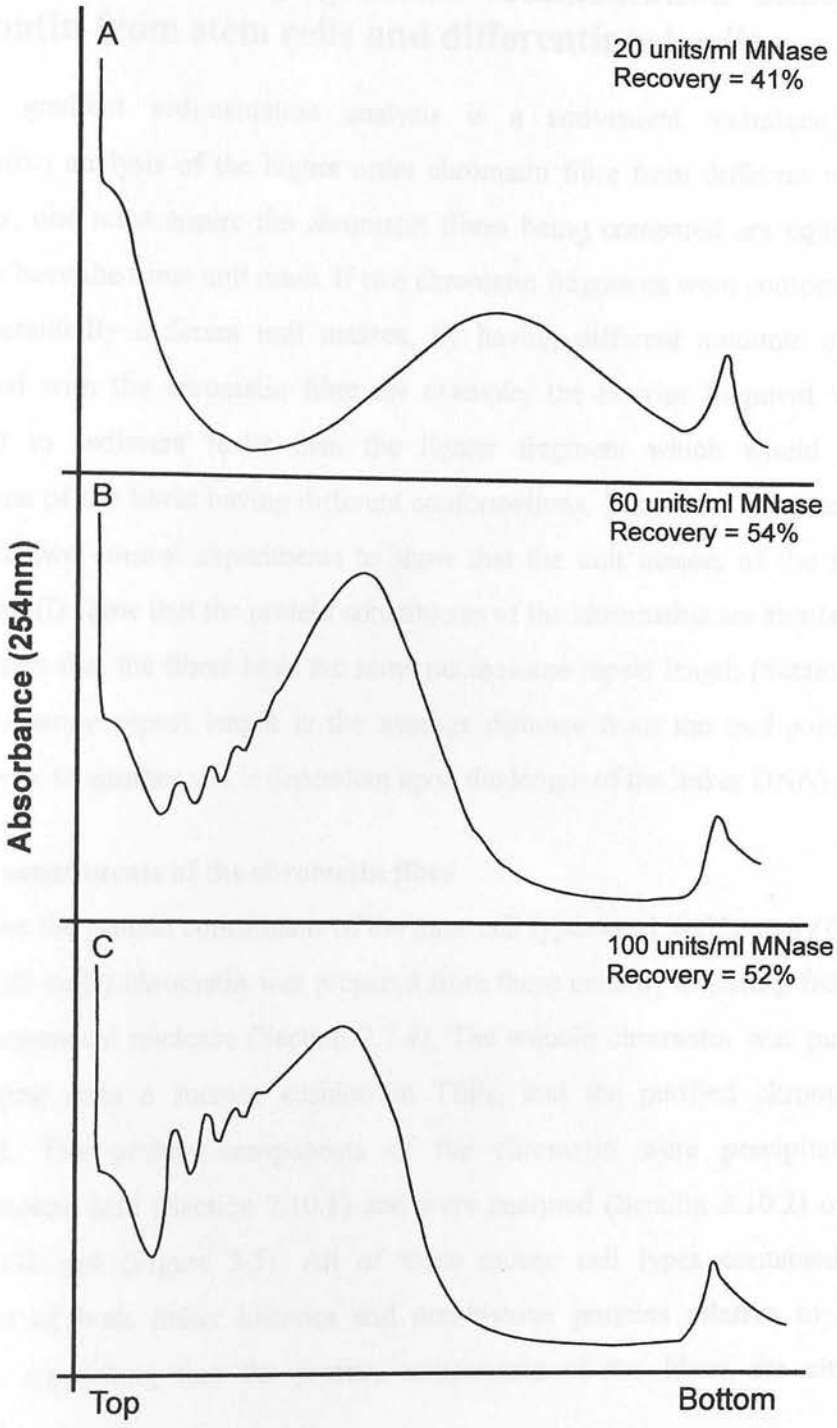


Figure 3.2 Sucrose gradient profiles of differently sized chromatin preparations. Nuclei were digested with different amounts of micrococcal nuclease for a set amount of time. Soluble chromatin was released overnight and the chromatin concentration was determined by spectroscopy. The chromatin samples were then fractionated on an isokinetic sucrose gradient.

3.3 Sucrose density gradient sedimentation analysis of chromatin from stem cells and differentiated cells

Sucrose gradient sedimentation analysis is a convenient technique for the comparative analysis of the higher order chromatin fibre from different cell types. However, one must ensure the chromatin fibres being compared are equivalent in that they have the same unit mass. If two chromatin fragments were compared which had substantially different unit masses, by having different amounts of protein associated with the chromatin fibre for example, the heavier fragment would be expected to sediment faster than the lighter fragment which would give the impression of the fibres having different conformations. Therefore, it is necessary to undertake two control experiments to show that the unit masses of the fibres are equivalent (i) show that the protein constituents of the chromatins are similar and (ii) demonstrate that the fibres have the same nucleosome repeat length (Section 1.1.10; the nucleosome repeat length is the average distance from the mid-point of one nucleosome to another and is dependent upon the length of the linker DNA).

Protein constituents of the chromatin fibre

To analyse the protein constituents of the three cell types used in this study (NIH3T3, F9 and ht2 cells) chromatin was prepared from these cells by digesting fresh nuclei with micrococcal nuclease (Section 2.7.4). The soluble chromatin was purified by centrifuging onto a sucrose cushion in TEP_{80} and the purified chromatin was collected. The protein components of the chromatin were precipitated with trichloroacetic acid (Section 2.10.1) and were analysed (Section 2.10.2) on a 15% SDS-PAGE gel (Figure 3.3). All of these mouse cell types contained similar quantities of both linker histones and non-histone proteins relative to the core histones, suggesting that the protein components of the fibres are similar, as expected.

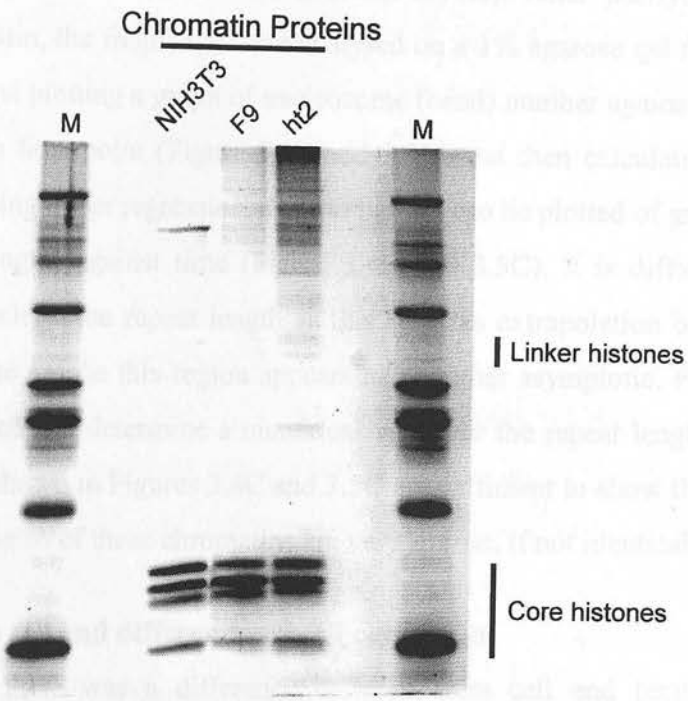


Figure 3.3 Protein constituents of NIH3T3, F9 and ht2 chromatin. Proteins were precipitated from purified chromatin by 20% TCA and were analysed on a 15% SDS-PAGE.

Repeat length of the chromatin fibre

Different nucleosome repeat lengths have been found in different cell types (Van Holde, 1988) for example Sea Urchin sperm has a repeat length of 260 bp (Spadafora et al., 1976) whilst HeLa cells have a repeat length of 190 bp. A chromatin fibre which has a shorter repeat length will have more nucleosomes per length of DNA than a chromatin fibre which has a longer repeat length. In turn, these differences in the nucleosome repeat length will affect the mass of the chromatin fibre and its sedimentation rate. To compare the nucleosome repeat lengths of NIH3T3, F9 and ht2 chromatins, nuclei prepared from these cell lines were digested with micrococcal nuclease across an exponential time course (Section 2.11.2). After purifying the DNA from the chromatin, the fragments were analysed on a 1% agarose gel (Figure 3.4A and 3.5A). By first plotting a graph of nucleosome (band) number against DNA fragment size for each time point (Figure 3.4B and 3.5B) and then calculating the gradient of the lines using linear regression, a second graph can be plotted of gradient (nucleosome repeat length) against time (Figure 3.4C and 3.5C). It is difficult to determine an exact nucleosome repeat length as this requires extrapolation back to time zero, and from the graphs this region appears to be rather asymptotic. For my purposes it is unnecessary to determine a numerical value for the repeat length and analysis of the curves shown in Figures 3.4C and 3.5C are sufficient to show that the nucleosomes repeat lengths of these chromatins are very similar, if not identical.

Conformation of stem cell and differentiated cell chromatin

To establish whether there was a difference between stem cell and terminally differentiated cell chromatin fibres, I compared bulk chromatin from F9 stem and NIH3T3 cells. Chromatin was prepared as described (Section 2.7.4); the two chromatin preparations were centrifuged in 6-40% isokinetic sucrose gradients in the same SW41 rotor to ensure the running conditions were equivalent. After sucrose gradient sedimentation the chromatin was pumped off the gradients (Figure 3.6A) and the DNA was purified and analysed by agarose gel electrophoresis (Section 2.8.1) on a 0.7% agarose gel in TPE buffer (Figure 3.6B). The agarose gel was scanned using a 473nm laser and a 580nm band-pass filter on a Fuji FLA-2000. The size of the bands was determined from the DNA size markers which were a 1 kb

DNA ladder, and a HindIII digest of λ, using Aida v2.0 analysis software. The DNA fragment size was plotted against the fraction number (Figure 3.6C) resulting in a graph relating the size of the fragment to its sedimentation rate which is indicative of the conformation of the chromatin fibres. The results show that the two chromatins have very similar sedimentation profiles suggesting that the conformation of the two chromatins are very similar.

Embryonic stem cells are unique as they have the potential to differentiate into all other mouse tissues (Section 3.1). The supposition is that these cells must have a chromatin structure which is poised to allow differentiation down all lineages, and therefore it is possible these cells have an atypical higher order chromatin fibre. To assess whether stem cells have a unique higher order conformation I compared terminally differentiated NIH3T3 chromatin to h2 stem cell chromatin. Soluble chromatin was prepared from the cells and sedimented on a sucrose gradient (Figure 3.7A), as for the previous experiment, and were analysed to identify any possible differences in the higher order chromatin fibre conformation (Figure 3.7B and 3.7C). As for F9 stem cell chromatin I was unable to find a difference in the relationship between DNA size and sedimentation for the two cell types. Higher-molecular weight chromatin fragments would be expected to accentuate any differences in the fibre conformation, but even for large fragments of 14 kb there were no differences between them. These two results (Figure 3.6C and 3.7C) suggest that there is no inherent difference between stem cell and terminally differentiated cell chromatin in terms of their higher order chromatin conformation.

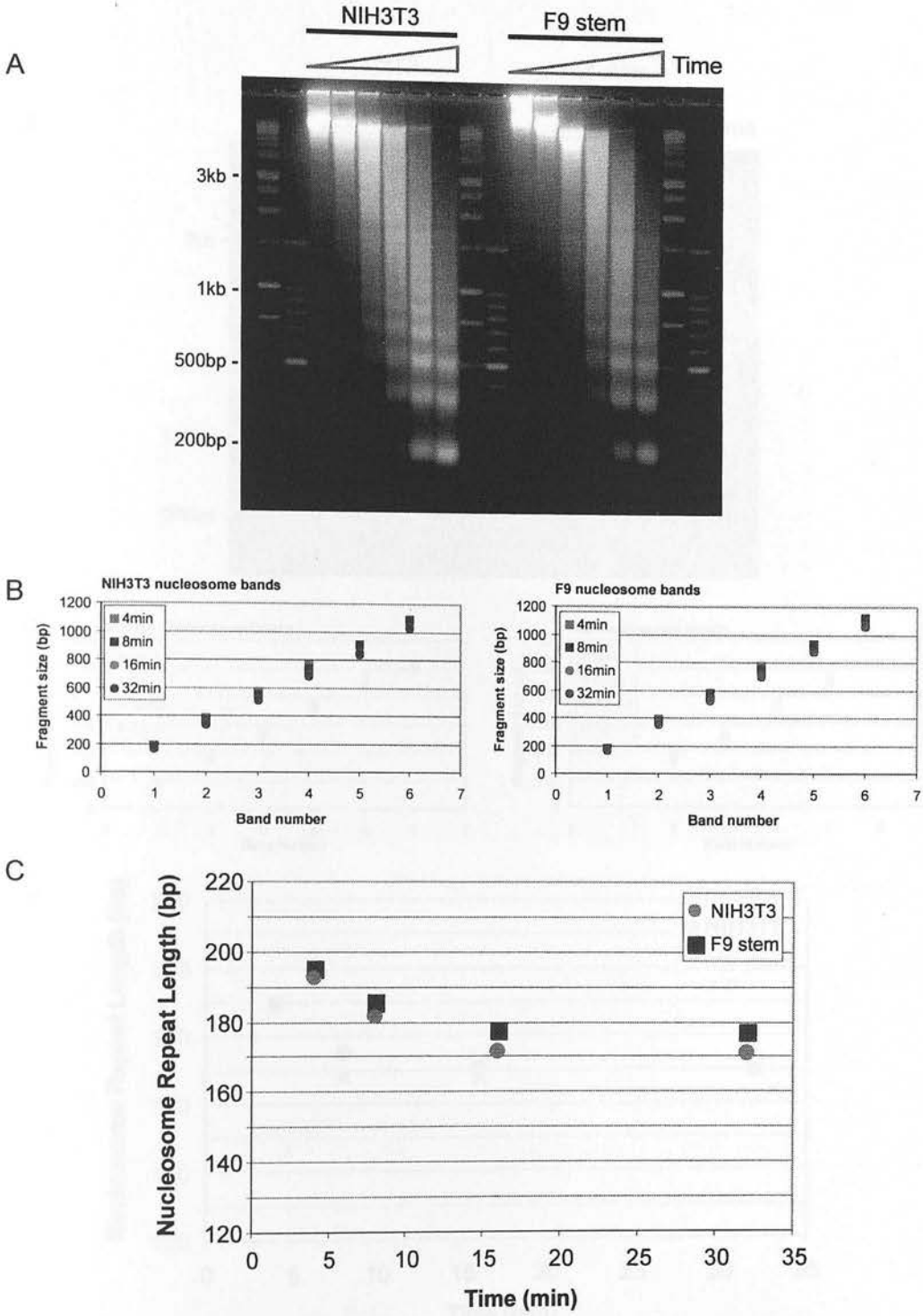


Figure 3.4 NIH3T3 and F9 cells have similar nucleosome repeat lengths. Nuclei were prepared from these cells and were digested with micrococcal nuclease for 1-32 minutes. The DNA was purified and analysed by agarose gel electrophoresis. The nucleosome repeat length was determined for the two cell types as described in the text. Markers are 1 kb and 100 bp ladder.

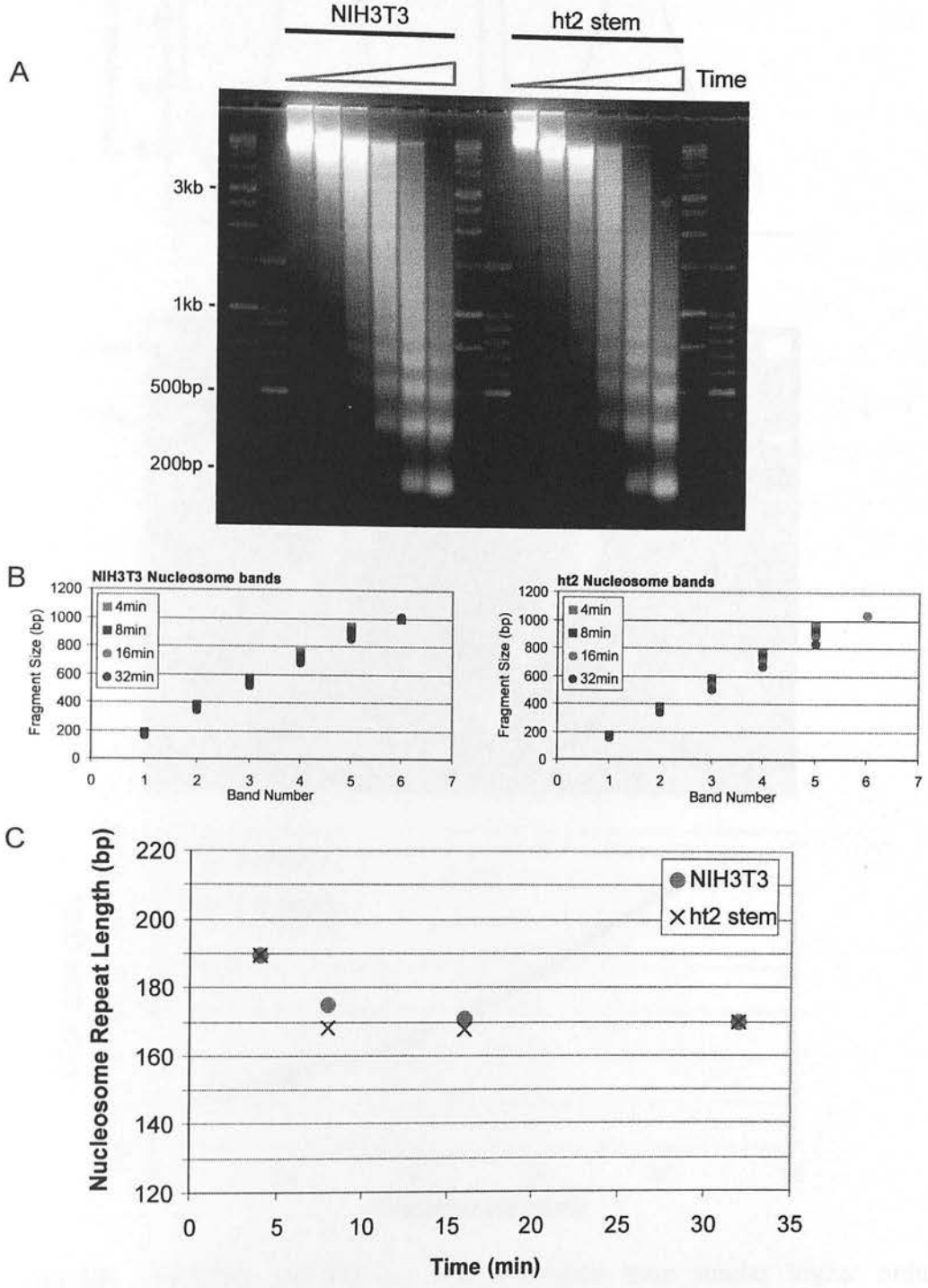


Figure 3.5 NIH3T3 and ht2 cells have similar nucleosome repeat lengths. Nuclei from cells were digested with micrococcal nuclease for 1-32 minutes. The DNA was purified and analysed by agarose gel electrophoresis. The nucleosome repeat length was determined for the two cell types as described in the text. Markers are 1 kb and 100 bp ladder.

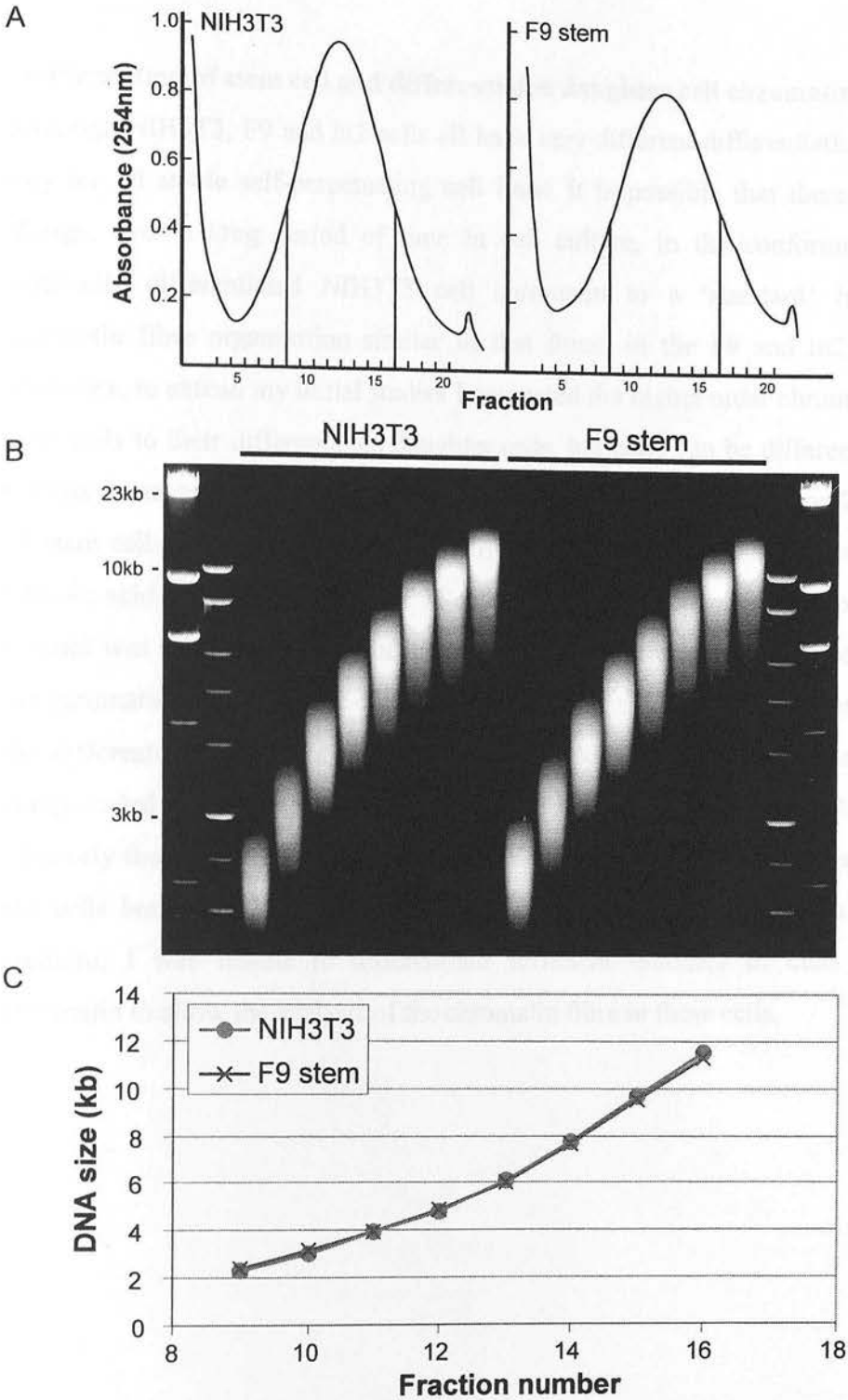


Figure 3.6 NIH3T3 and F9 cell bulk chromatin have similar higher order conformations. Soluble chromatin was prepared from the cells and analysed by sucrose gradient sedimentation. Chromatin fractions were isolated and the DNA was purified. Analysis of the DNA fragment sizes shows that the two cells types have very similar relationships between DNA size and the extent of gradient sedimentation indicating they have the same higher order chromatin fibre conformation. Markers are 1 kb ladder and λ-HindIII.

Conformation of stem cell and differentiated daughter cell chromatin

Although NIH3T3, F9 and ht2 cells all have very different differentiation potentials they are all stable self-perpetuating cell lines. It is possible that there has been a change, over a long period of time in cell culture, in the conformation of the terminally differentiated NIH3T3 cell chromatin to a 'standard' higher order chromatin fibre organisation similar to that found in the F9 and ht2 stem cells. Therefore, to extend my initial studies I compared the higher order chromatin fibre of stem cells to their differentiated daughter cells. ht2 cells can be differentiated by 3-methoxybenzamide (MBA) to give an epidermal-like cell type (Section 2.6.5) whilst F9 stem cells can be differentiated to parietal endoderm-like cells by the addition of retinoic acid and di-butyryl cAMP (Section 2.6.6). In a number of experiments an attempt was made to prepare sufficient amounts of differentiated ht2 cells to allow the chromatin structure to be analysed by sucrose gradient sedimentation. However, the differentiated ht2 cells are large, and tend to grow at a low density. This is compounded by MBA being fairly toxic (Figure 2.2) to these cells such that relatively few differentiated cells survived. Even after repeating the differentiation of ht2 cells both with MBA, and alternatively by withdrawing LIF from the culture medium, I was unable to differentiate sufficient numbers of cells to prepare chromatin to allow the analysis of the chromatin fibre in these cells.

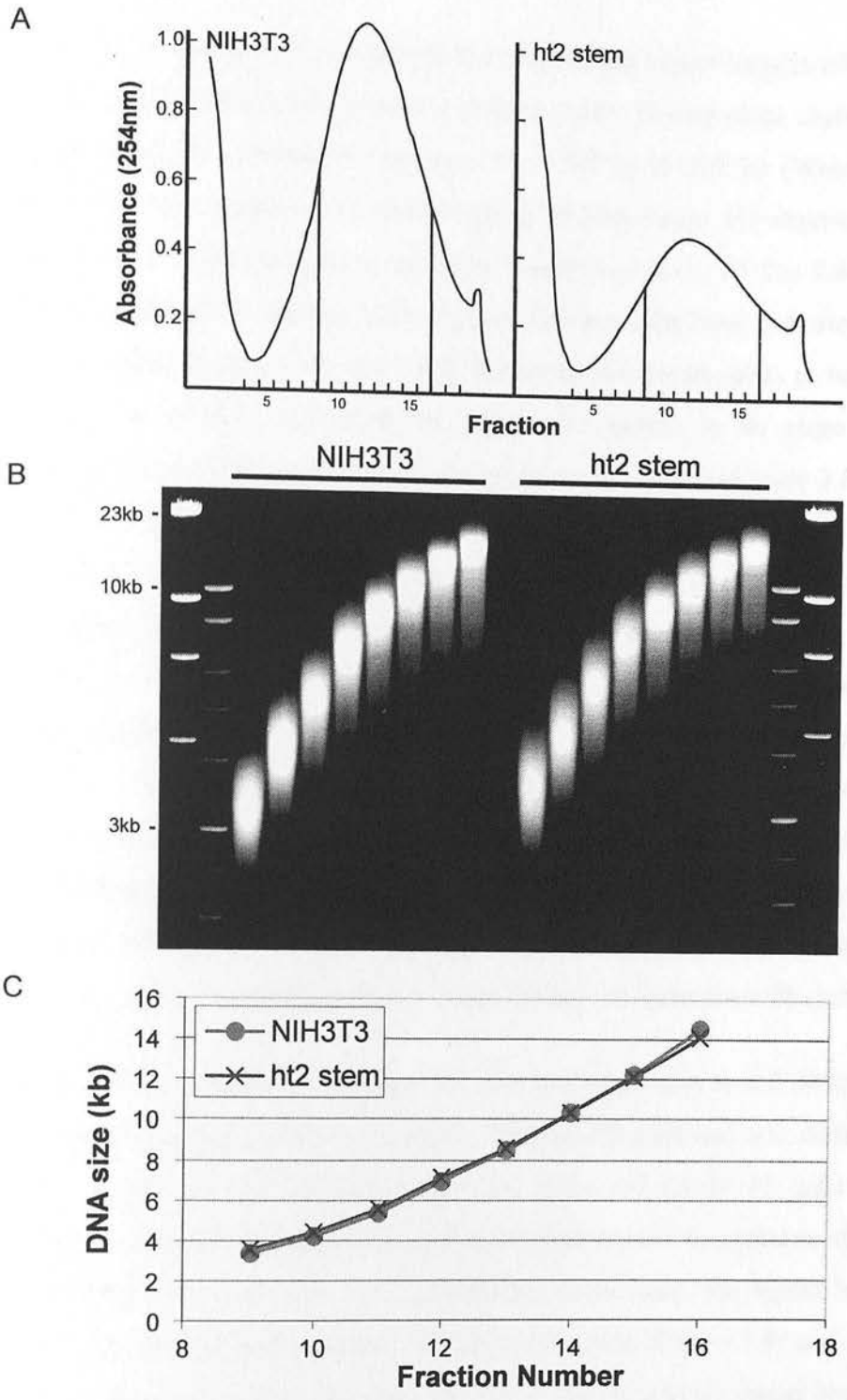


Figure 3.7 NIH3T3 bulk chromatin has a similar higher order conformation to ht2 stem cell chromatin. Soluble chromatin was prepared from the cells and analysed by sucrose gradient sedimentation. Chromatin fractions were isolated and the DNA was purified. Analysis of the DNA shows that the two cell types have the same relationship between DNA size and the extent of sedimentation suggesting they have the same higher order chromatin fibre conformation. Markers are 1 kb ladder and λ-HindIII.

However, it was possible to analyse the nucleosome repeat lengths of ht2 stem cell and h2 differentiated cell chromatin. (Figure 3.8A). During chick erythropoiesis the repeat length in erythrocytes increases from 190 bp to 212 bp (Weintraub, 1978). Although this corresponds to an increase in H5 expression, H5 expression has been shown not to alter the nucleosome repeat length (Sun et al., 1990c). Likewise for my results, preliminary analysis indicates that the stem cells have a shorter nucleosome repeat length (Figure 3.8B and 3.8C). However, this result needs to be treated with caution as in this experiment the stem cells appear to be more sensitive to micrococcal nuclease digestion than the differentiated cells (Figure 3.8A), probably due to the permeability of the nuclei to micrococcal nuclease and the initial rate of digestion for the stem cells is greater than for the differentiated cells. If the stem cells are digested more rapidly, it would give the impression of the repeat length being shorter due to trimming by the nuclease. Ideally, this analysis should be repeated using isolated chromatin, although as discussed above it was not possible to produce sufficient cells to isolate the amount of chromatin needed for this analysis. Further analysis indicates that in this experiment the stem cells are digested 3-fold faster than the differentiated cells so replotting the points shows that the differences seen in Figure 3.8C are less than they appear, and are in contrast to previous studies for the alteration of the nucleosome repeat length during cell maturation (Weintraub, 1978).

As it was not possible to compare ht2 stem cell chromatin to ht2 differentiated cell chromatin directly I studied chromatin from undifferentiated and differentiated F9 cells. In the presence of retinoic acid and di-butyryl cAMP F9 cells differentiate rapidly to parietal endoderm-like cells and therefore provide a good model system for the differentiation of cells. As a preliminary experiment, the repeat lengths of F9 stem cells and F9 differentiated cells were compared (Figure 3.9) and it was found that, as for the ht2 cells, the stem cell has a slightly shorter repeat length than the differentiated cell although this is probably insufficient to affect the hydrodynamic gradient analysis. Soluble chromatin was prepared from F9 stem cells and differentiated F9 cells and the chromatins were analysed by sucrose gradient sedimentation (Figure 3.10A). The chromatin was purified and analysed on a 1% agarose gel in TBE buffer (Figure 3.10B) and the size of the bands was determined

and plotted against the fraction numbers (Figure 3.1C). The points have been offset as the two gradients were pumped off the gradient out of synchronisation with each other. This does not alter the results as if there was a difference between the conformation of the two chromatins the gradients of lines would be different, as any differences would be accentuated for larger fragments. Therefore, by this hydrodynamic approach I was unable to find any changes in the conformation of the higher order chromatin fibre between undifferentiated stem cells and differentiated cells.

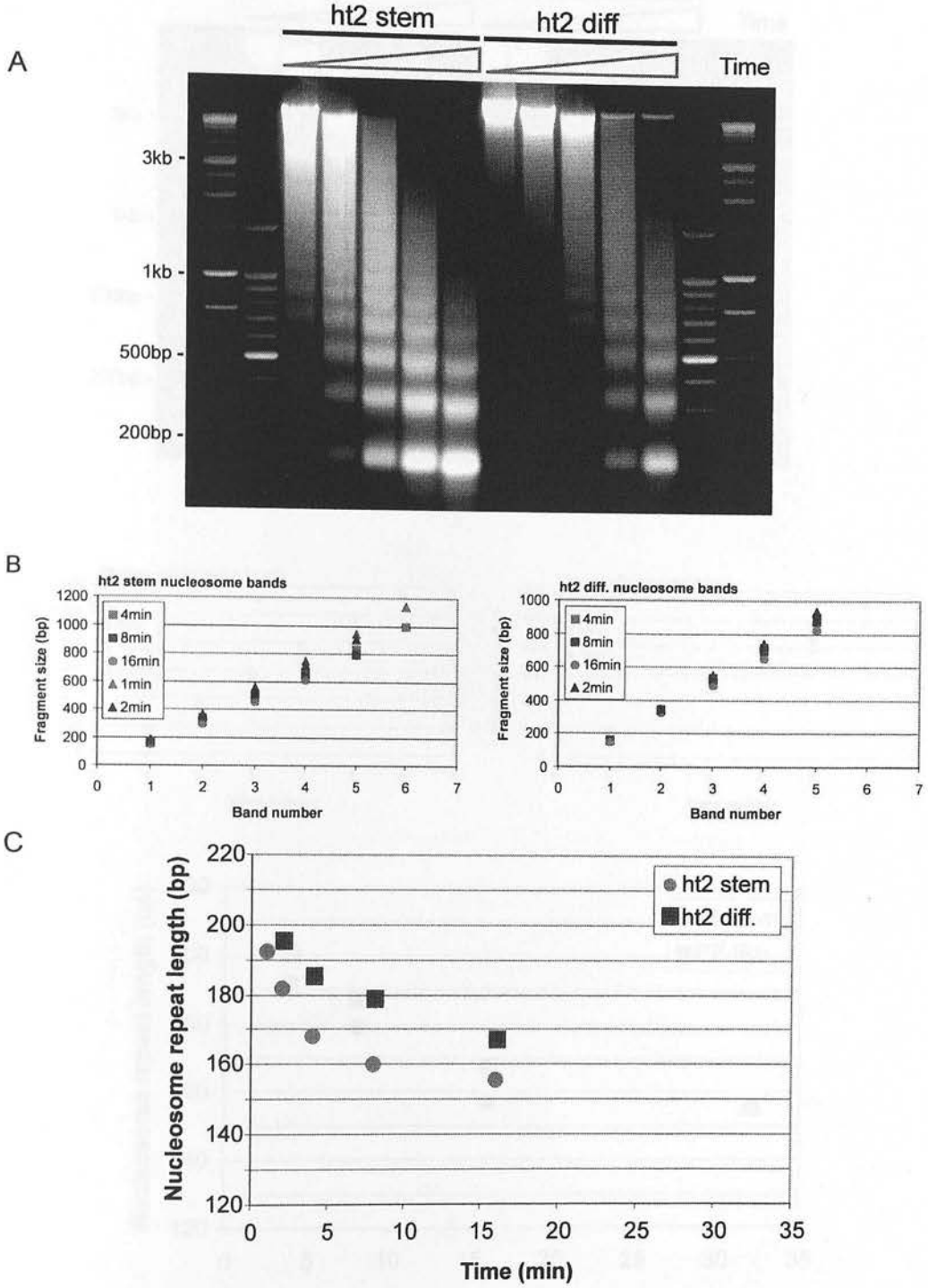


Figure 3.8 The nucleosome repeat lengths of ht2 stem cells and differentiated ht2 cells are similar. Nuclei prepared from cells were digested with micrococcal nuclease for 1-16 minutes. The DNA was purified and analysed by agarose gel electrophoresis. The nucleosome repeat length was determined for the two cell types as described in the text. Markers are 1 kb and 100 bp ladder.

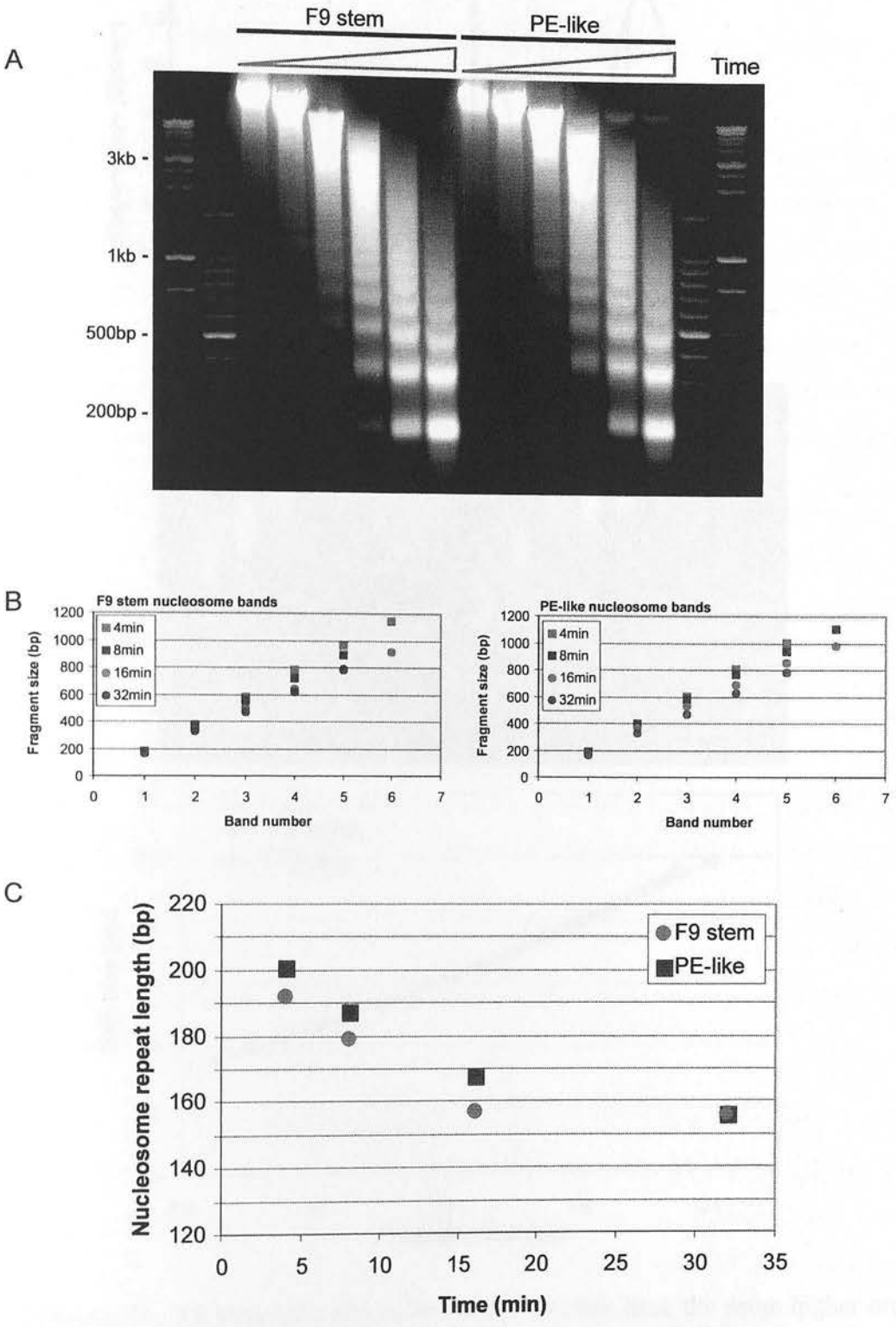


Figure 3.9 The nucleosome repeat lengths of F9 stem cells and differentiated F9 cells are similar. Nuclei from cells were digested with micrococcal nuclease for 1-32 minutes. The DNA was purified and analysed by agarose gel electrophoresis. The nucleosome repeat length was determined for the two cell types as described in the text. Markers are 1 kb and 100 bp ladder.

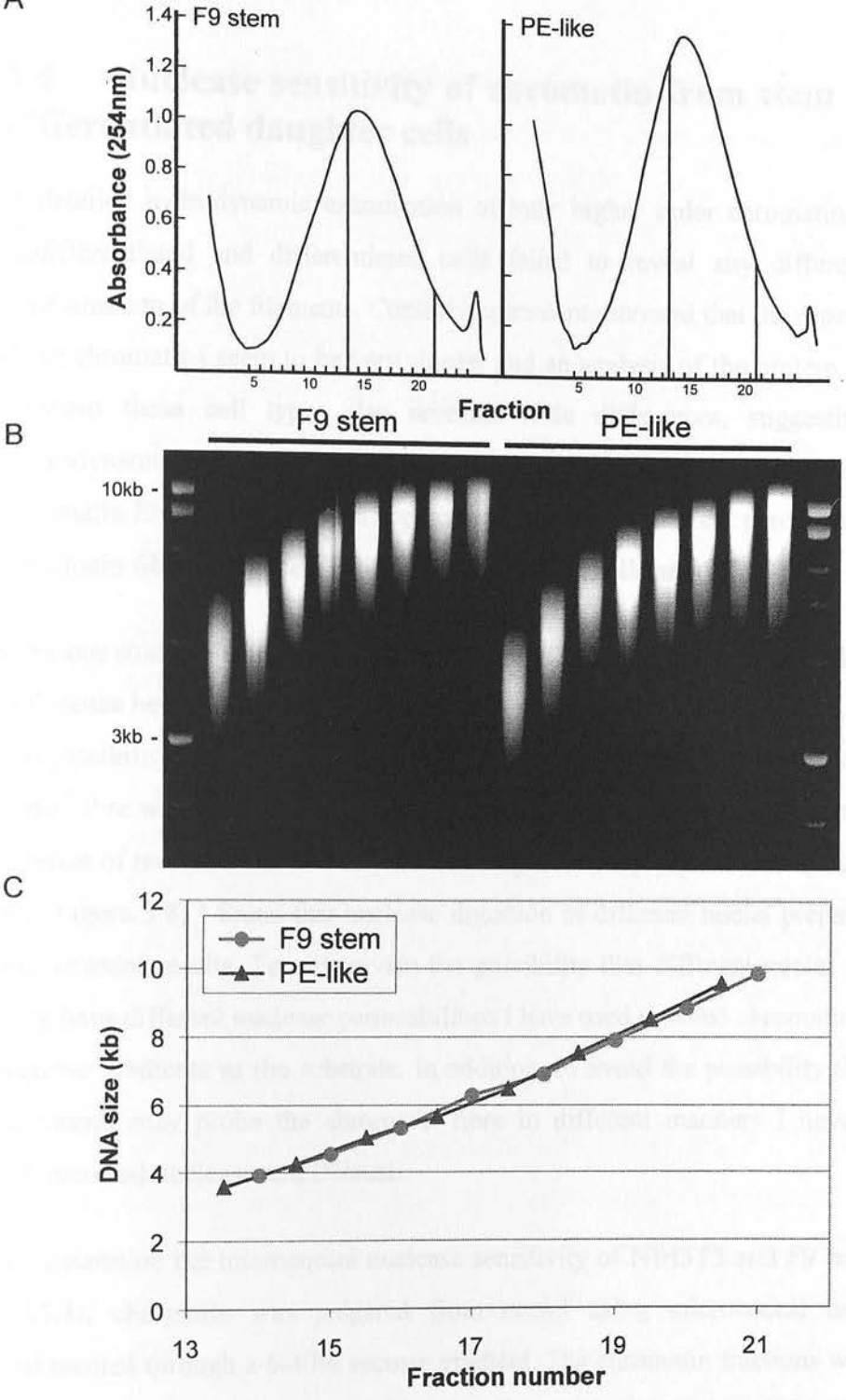


Figure 3.10 F9 stem cells and differentiated F9 cells have the same higher order chromatin conformation. Soluble chromatin was prepared from the cells and analysed by sucrose gradient sedimentation. Chromatin fractions were isolated and the DNA was purified. Analysis of the DNA shows that the two cell types have the same relationship between DNA size and the extent of sedimentation suggesting they have the same higher order chromatin fibre conformation. Markers are 1 kb ladder and λ-HindIII.

3.4 Nuclease sensitivity of chromatin from stem cells and differentiated daughter cells

A detailed hydrodynamic examination of bulk higher order chromatin fibres from undifferentiated and differentiated cells failed to reveal any differences in the conformation of the filaments. Control experiments showed that the repeat lengths of these chromatins seem to be very similar and an analysis of the protein composition between these cell types also revealed little differences, suggesting that the hydrodynamic studies are probably an accurate reflection of the conformation of the chromatin fibres. These results appear surprising as the current perception is that the chromatin fibres from stem cells and differentiated cells are different.

Previous studies using nucleases (Huebner et al., 1981) have suggested that there is a difference between the higher order fibre of stem and differentiated cells. To explore the possibility that there may indeed be differences in the conformation of the higher order fibre which were undetected using a hydrodynamic approach, I also undertook a series of nuclease accessibility studies. In preliminary experiments (data not shown and Figure 3.8) I found that nuclease digestion of different nuclei preparations gave inconsistent results. To circumvent the possibility that different nuclei preparations may have different nuclease permeabilities I have used purified chromatin taken from sucrose gradients as the substrate. In addition, to avoid the possibility that different nucleases may probe the chromatin fibre in different manners I have used both micrococcal nuclease and DNaseI.

To determine the micrococcal nuclease sensitivity of NIH3T3 and F9 cells (Section 2.11.4), chromatin was prepared from nuclei using micrococcal nuclease and sedimented through a 6-40% sucrose gradient. The chromatin fractions were pumped off the gradients (Figure 3.11A) and stored in sucrose TEP_{80} at 4°C until analysed. Equivalently sized peak fractions were taken from each gradient and the concentration of the chromatins was equalised using a suitable concentration of sucrose in TEP_{80}. Control experiments were undertaken (data not shown) where the chromatin concentration was not adjusted, but different amounts of nuclease were added to the samples, and the results were identical to those where the chromatin

concentrations were adjusted. In addition all of the chromatin samples tested negative for endogenous nucleases within the time-frame of the experiments. The chromatin samples were supplemented with calcium chloride to 3 mM and an aliquot was removed into micrococcal nuclease stop buffer as time zero. After the start of digestion with micrococcal nuclease aliquots of the reaction were removed at exponential time intervals into stop buffer. The DNA was purified from the samples and analysed on an agarose gel (Figure 3.11B). The results show that there is little difference between the digestion rates of the NIH3T3 and stem cells. Possibly, the NIH3T3 cells are fractionally more sensitive to the nuclease although this distinction might be beyond the resolution of the experiment.

If there was a difference between the higher order chromatin fibres between stem cells and differentiated cells it would be expected to be accentuated for pluripotential embryonic stem cells, as these cells would be expected to be primed in most developmental lineages and to only repress these lineages after differentiation. To compare the sensitivity of the ht2 cells to NIH3T3 cells a similar experiment was performed to the one described above. Purified chromatin was fractionated from both NIH3T3 and ht2 cells and digested under controlled conditions using micrococcal nuclease as described above. The DNA was purified from the samples and analysed on a 1.2% TBE gel (Figure 3.12). It is clearly apparent that ht2 ES cells are less sensitive to micrococcal nuclease than the terminally differentiated NIH3T3 cells. As the NIH3T3 cells are digested about four-fold faster than the ht2 chromatin it suggests that the pluripotential ht2 cells are less sensitive to micrococcal nuclease than the totipotential F9 cells. These experiments were repeated with both independent chromatin gradients and with different pairs of fractions from each gradient giving essentially the same results, suggesting that this is a specific difference between pluripotential stem cells and terminally differentiated NIH3T3 cells.

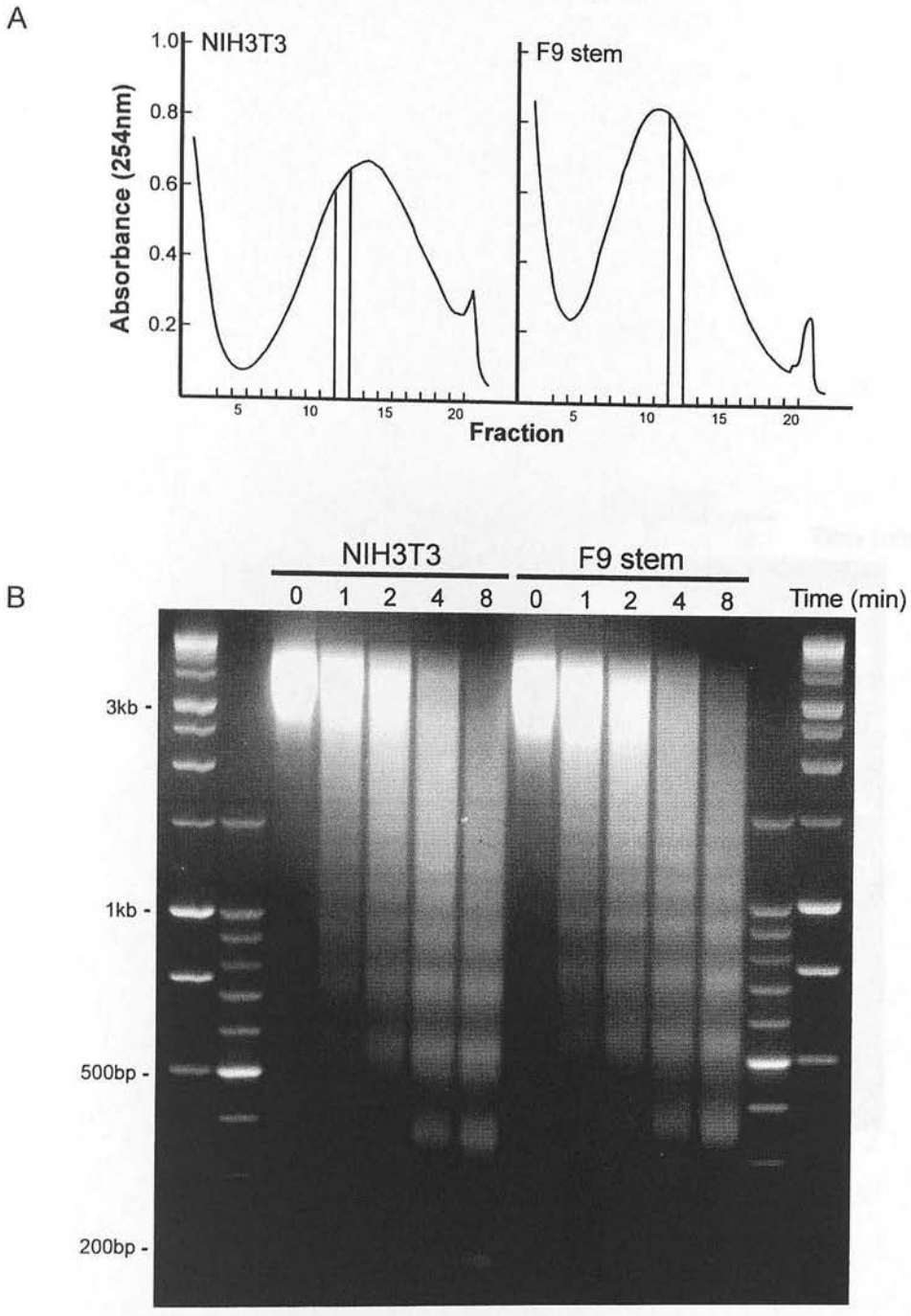


Figure 3.11 NIH3T3 chromatin is similarly sensitive to micrococcal nuclease as F9 stem cell chromatin. Purified NIH3T3 and F9 cell chromatin were diluted to the same concentration in TEP_{80}-sucrose. 16 units/ml micrococcal nuclease was added and aliquots were removed over 8 minutes. The DNA was purified and analysed by agarose gel electrophoresis. Markers are 1 kb and 100 bp ladders.

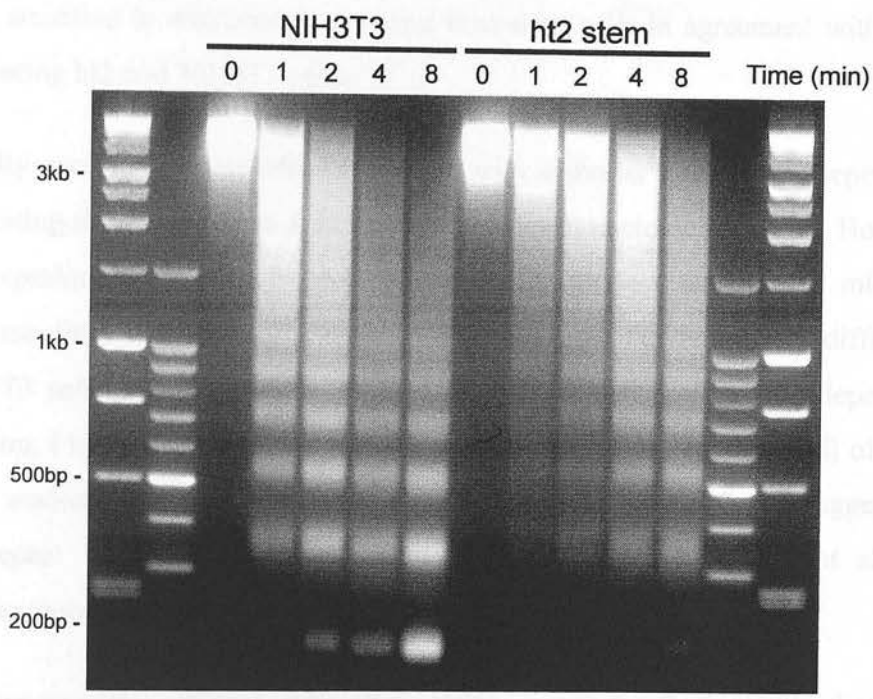


Figure 3.12 ht2 stem cell chromatin is less sensitive to micrococcal nuclease than chromatin from terminally differentiated cells. Purified NIH3T3 and ht2 stem cell chromatin were diluted to the same concentration in TEP_{80}-sucrose. 10 units/ml micrococcal nuclease was added and aliquots were removed over 8 minutes. The DNA was purified and analysed by agarose gel electrophoresis. Markers are 1 kb and 100 bp ladders.

NIH3T3 cells have been selected over many years to replicate relatively rapidly in culture which may have facilitated an alteration in the chromatin structure, such than they are no longer representative of the terminally differentiated state. Due to technical restrictions, as mentioned previously, I found it impossible to produce sufficient differentiated ht2 cells for chromatin analysis. Therefore I decided to compare the nuclease sensitivity of F9 stem cells and differentiated F9 cells. Chromatin was prepared from these cells using micrococcal nuclease and purified on sucrose gradients as for the previous experiments. Peak fractions were taken from these gradients and digested with micrococcal nuclease over a time course. Analysis of the chromatin (Figure 3.13) shows that the differentiated F9 cells are substantially more sensitive to micrococcal nuclease than stem cells in agreement with my data comparing ht2 and NIH3T3 cells.

Rapidly cycling cells are often associated with a shorter nucleosome repeat length, suggesting they might have a different sensitivity to nuclease digestion. However, in my experiments, the rapidly cycling stem cells are less sensitive to micrococcal nuclease than both the differentiated F9 cells and the terminally differentiated NIH3T3 cells, suggesting that these results are not cell proliferation dependent. In addition, I have already shown that the nucleosome repeat lengths of all of the cells being studied are approximately the same (Figure 3.4, 3.5, 3.6, 3.17), suggesting that the repeat lengths of the cells used in these experiments do not show any proliferation dependence.

Another possible, although difficult to define, reason for these results showing that differentiated cells are more sensitive to nucleases is that micrococcal nuclease has some specificity for differentiated cell chromatin. To investigate this I repeated the previous experiment but digested the chromatin with DNaseI instead of micrococcal nuclease. Purified chromatin was fractionated from F9 stem cells and differentiated F9 cells and equivalent fractions were taken from each gradient and adjusted to the same concentration of chromatin (for one experiment the chromatin samples were not diluted, but digested with different amounts of DNaseI). Samples were supplemented with magnesium and calcium ions and an aliquot was removed into

stop buffer. To the remaining material, a small amount of DNaseI was added to each sample and aliquots were removed over a period of 8 minutes. Purification of the DNA and analysis by agarose gel electrophoresis (Figure 3.14) shows that DNaseI, like micrococcal nuclease, digests differentiated cell chromatin more rapidly than stem cell chromatin. Repeating this experiment with independent chromatin preparations gave the same result, as did taking adjacent fractions, and digesting with micrococcal nuclease.

These results clearly show that although there is no difference in the sedimentation profile of chromatin from differentiated cells and stem cells there is a clear and reproducible difference in the sensitivities of these chromatins to nucleases. This suggests that the altered sensitivities to nucleases of the chromatin fibre is not a consequence of a simple difference in the compaction of the chromatin fibre, but reflects a more subtle conformational change which cannot be identified by sucrose gradient sedimentation analysis.

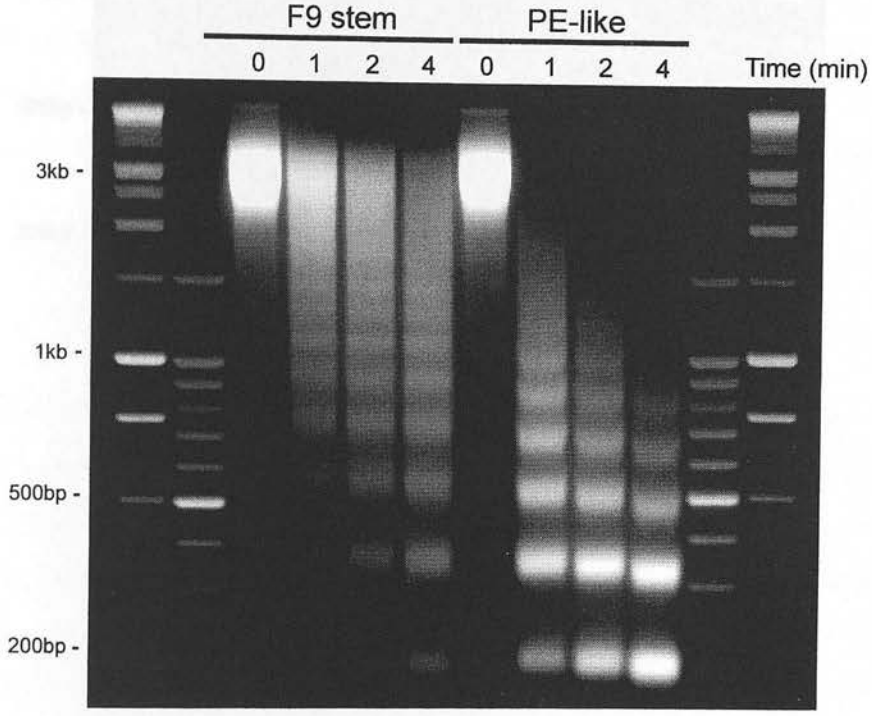


Figure 3.13 Differentiated F9 cell chromatin is more sensitive to micrococcal nuclease digestion than that of F9 stem cells. Purified F9 and differentiated F9 cell chromatin were diluted to the same concentration in TEP_{80}-sucrose. 10 units/ml micrococcal nuclease was added and aliquots were removed over 4 minutes. The DNA was purified and analysed by agarose gel electrophoresis. Markers are 1 kb and 100 bp ladders.

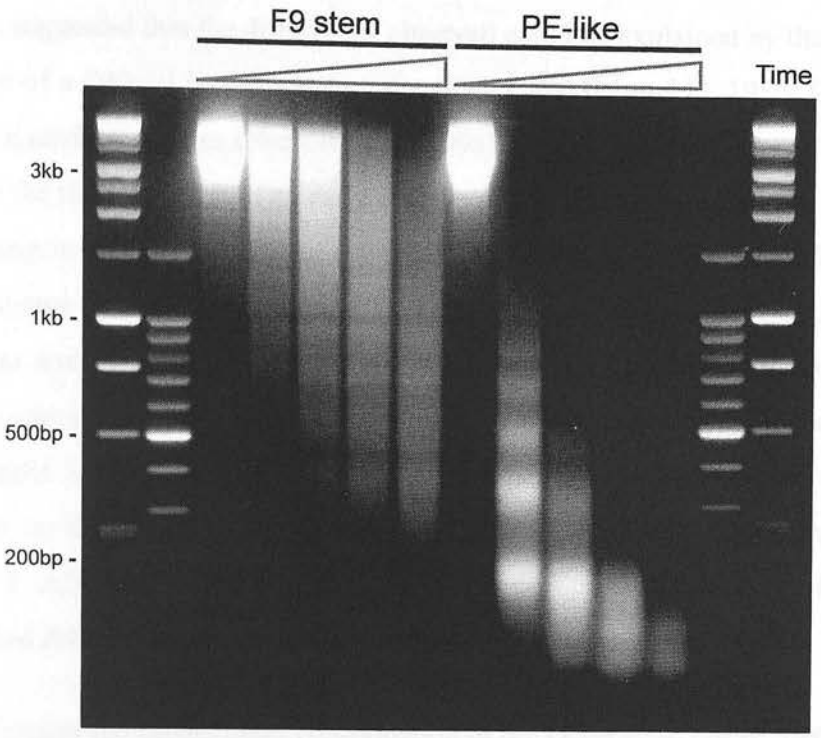


Figure 3.14 Differentiated cell chromatin is more sensitive to DNaseI than stem cell chromatin. Purified F9 and differentiated F9 cell chromatin were diluted to the same concentration in TEP_{80}-sucrose. 2 units/ml DNaseI was added and aliquots were removed over 8 minutes. The DNA was purified and analysed by agarose gel electrophoresis. Markers are 1 kb and 100 bp ladders.

3.5 Conformation of the higher order chromatin fibre around the Oct-4 gene in ES cells

Originally, it was thought that the active β-globin loci had a more open conformation than the inactive ovalbumin loci (Kimura et al., 1983), although later more detailed studies suggested that the differences observed could be explained by the presence or absence of a DNaseI hyper-sensitive site (Fisher and Felsenfeld, 1986; Caplan et al., 1987). Knockout studies (Nichols et al., 1998) have shown that Oct-4 is an essential protein for the formation of an early embryo and in its absence the inner cell mass is not pluripotent. As current data suggest that Oct-4 is essential for maintaining the pluripotency of stem cells, I thought it would be interesting to determine whether this gene has a unique higher order chromatin conformation in embryonic stem cells that could reflect its critical role, even though in our lab we have never found any differences in the conformation of active and inactive loci. Although I have shown there is no difference in the bulk chromatin structure between ht2 stem cells and NIH3T3 cells (Figure 3.7C), the hydrodynamic approach used would not have identified differences at specific chromosomal loci.

To determine the higher order conformation of the Oct-4 gene I have used a variation on our hydrodynamic approach which is described in more detail in Chapter 4. The sucrose gradients described so far have all been sedimented in TEP_{80} which will maintain the chromatin in a folded state. Due to this, these TEP_{80} gradients are separating the chromatin based on both its size and higher order conformation. Therefore, within one isokinetic fraction there could be two populations of chromatin of different sizes if they have different higher order conformations. By dialysing an isokinetic fraction into TEP_5 the higher order conformation of the chromatin fibre is lost, such that when the sample is run on a TEP_5 sucrose gradient the chromatin will be fractionated on size rather than higher order conformation. Therefore, if a specific chromosomal loci has a different chromatin conformation it will be separated away from the bulk chromatin in the second TEP_5 sucrose gradient.

ht2 chromatin was prepared from nuclei and sedimented on a TEP_{80} sucrose gradient (Figure 3.15A) as described previously. Fraction 14 was taken and dialysed into

TEP_5 in a micro-dialyser (Gibco), concentrated in a micro-concentrator (Amicon), and then sedimented through a 6-40% sucrose gradient in TEP_5 in an SW41 rotor for 3 hours at 48,000rpm. Fractions 9-16 were taken from the TEP_5 gradient (Figure 3.15B) and the DNA was purified and dot blotted onto a charged nylon membrane (Section 2.8.4). In addition a serial dilution of sonicated ES cell chromatin was dot blotted on to the membrane as a transfer and hybridisation control.

To identify the Oct-4 gene a 680 bp StuI probe encompassing the POU domain was prepared from an Oct-4 genomic clone (Plasmid #89) and was tested for its specificity by hybridising to an EcoRI or BamHI mouse cell genomic DNA blot (Section 2.8.3). The results (Figure 3.16) show that the probe hybridises specifically and is therefore suitable for analysing the chromatin conformation in this region.

To identify whether the chromatin structure of the Oct-4 locus was different to bulk chromatin the dot blot was probed with the StuI probe (Figure 3.17A) and the membrane was exposed on a phosphor screen and scanned on a phosphorimager (Fuji, FLA200). The blot was subsequently probed with a B2 interspersed repeat (Krayev et al., 1982) derived from the murine LIF receptor (Chambers et al., 1997; Plasmid #84) which will hybridise to the bulk chromatin. Once again the membrane was exposed to a phosphor screen and scanned on a phosphorimager. The amount of signal present in each dot was determined and from the serial dilution of control DNA a non-linear calibration curve was produced which allowed the signal from the two sample blots to be equalised. The equalised signals are plotted against fraction number (Figure 3.17B). From the graph the Oct-4 peak coincides directly with the bulk chromatin peak suggesting that the conformation of the higher order fibre around this critical gene is equivalent to the bulk chromatin fibres in agreement with earlier studies for the chicken β-globin and ovalbumin genes (Fisher and Felsenfeld, 1986; Caplan et al., 1987). This experiment was repeated for another fraction giving identical results.

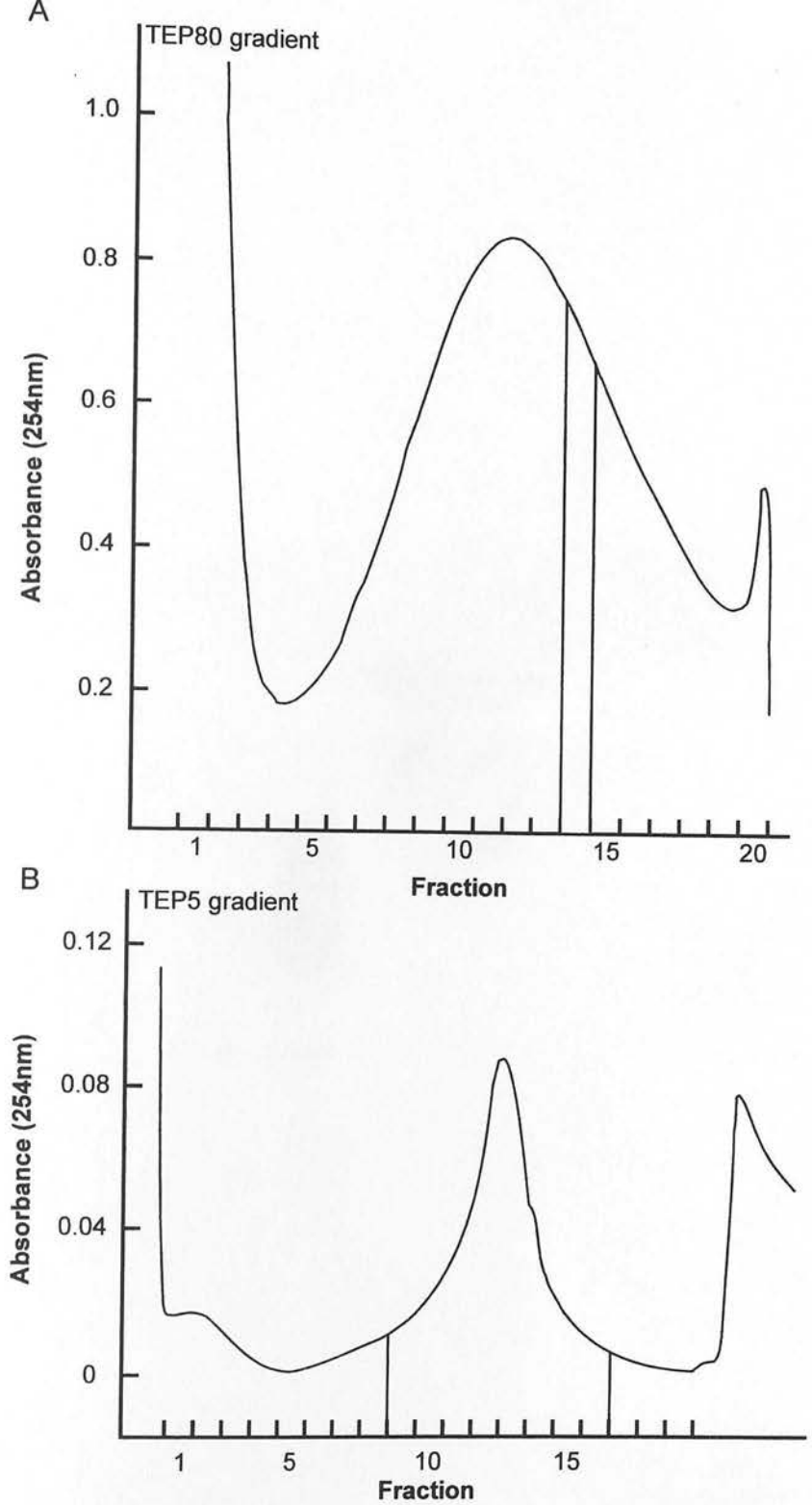


Figure 3.15 ht2 chromatin fractionated by size and conformation in a TEP_{80} sucrose gradient, and by size alone in a TEP_5 sucrose gradient. (A) Soluble chromatin was isolated from ht2 cells and fractionated on a sucrose gradient in TEP_{80}. (B) An isokinetic fraction was taken and dialysed into TEP5 and subsequently size fractionated on a sucrose gradient in TEP_5.

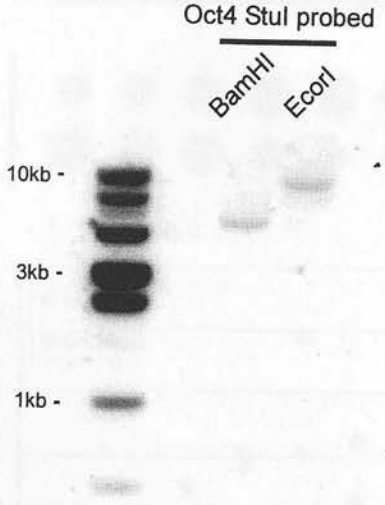


Figure 3.16 A StuI probe derived from the Oct-4 locus is specific. A StuI fragment encompassing the POU-domain of Oct-4 was used to probe a mouse genomic BamHI and EcoRI blot. The presence of a single band indicates the probe is specific and does not cross-react to other parts of the genome.

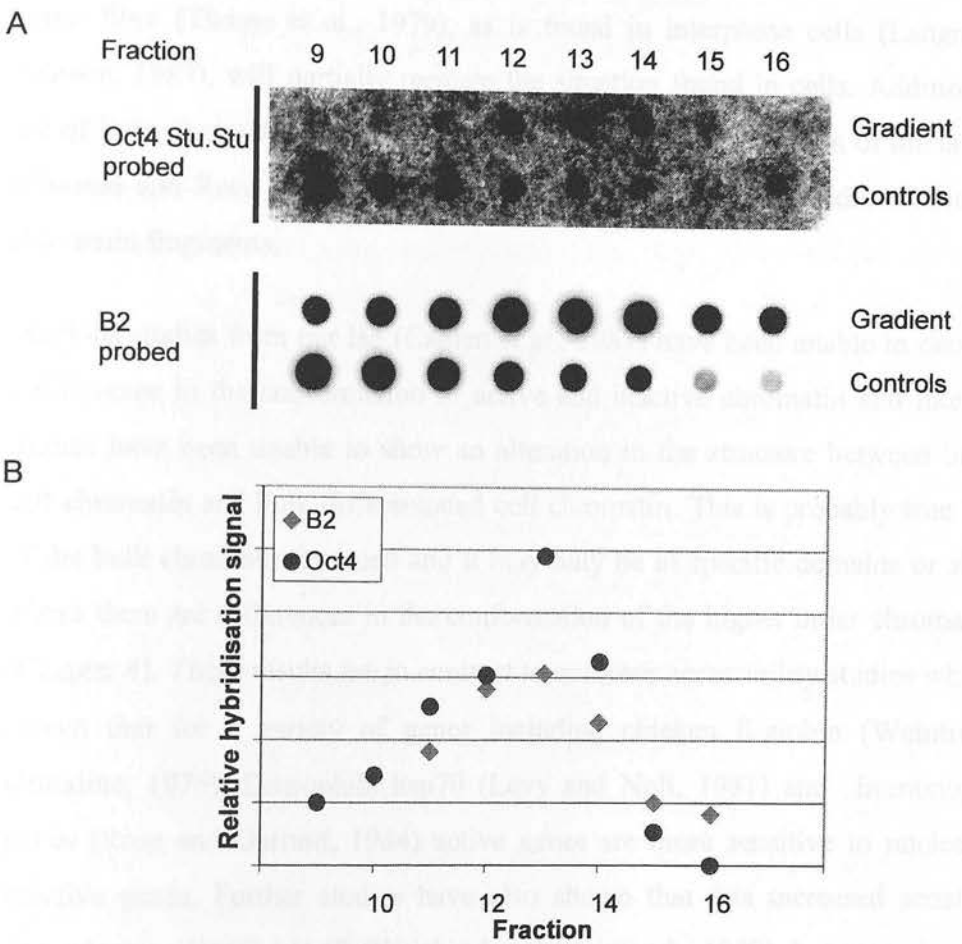


Figure 3.17 In ht2 ES cells the Oct-4 locus has the same higher order chromatin conformation as bulk fibres. Chromatin from ht2 ES cells was fractionated in a sucrose gradient containing TEP_{80}. An isokinetic aliquot was dialysed into TEP_5 and fractionated in a sucrose gradient in TEP_5. (A) Purified DNA from the TEP_5 gradient was dot-blotted with a probe to the Oct-4 locus to determine the peak position of the Oct-4 chromatin. (B) The blot was re-hybridised with a probe to a B2 repeat to determine the position of the peak bulk chromatin. (C) The hybridisation signals were quantified from the control dots and were equalised and plotted against fraction number.

3.6 Discussion

In this series of experiments I have examined whether the conformation of the chromatin fibre changes during the differentiation of cells. The approach I have chosen to use can only give a partial picture of the conformation of the chromatin fibre within cells as the influences on the chromatin structure by torsional constraints or attachment to the nuclear matrix have been lost. However, by maintaining the chromatin fibres in an ionic environment in which the chromatin will be folded into a 30nm fibre (Thoma et al., 1979), as is found in interphase cells (Langmore and Paulson, 1983), will partially recreate the situation found in cells. Additionally the use of large chromatin fragments will ensure that the redistribution of linker histones (Thomas and Rees, 1983) will not substantially influence the sedimentation of the chromatin fragments.

Previous studies from our lab (Caplan et al., 1987) have been unable to demonstrate a difference in the conformation of active and inactive chromatin and likewise my studies have been unable to show an alteration in the structure between bulk stem cell chromatin and bulk differentiated cell chromatin. This is probably true for most of the bulk chromatin in a cell and it may only be in specific domains or situations where there are differences in the conformation of the higher order chromatin fibre (Chapter 4). These results are in contrast to nuclease accessibility studies which have shown that for a variety of genes including chicken β-globin (Weintraub and Groudine, 1976), Drosophila hsp70 (Levy and Noll, 1981) and immunoglobulin genes (Rose and Garrard, 1984) active genes are more sensitive to nucleases that inactive genes. Further studies have also shown that this increased sensitivity is dependent on HMG-14/-17 (Weisbrod and Weintraub, 1979), but more importantly these results probably just reflect a partial loss or replacement of histone H1 and the acetylation of core histones which might alter the stability of the chromatin fibre.

In my experiments both micrococcal nuclease and DNaseI reproducibly showed that differentiated cell chromatin was more sensitive to nuclease digestion than stem cell chromatin. This is in contrast to the widely held belief that pluripotential stem cells must have a more 'open' chromatin structure to allow them to undergo a pathway of

differentiation. Previous studies by Huebner (1981) have also suggested that stem cell chromatin is more sensitive to nuclease digestion than differentiated cell chromatin. It is difficult to reconcile my data to this previous data, except to say that the reproducible nature of my studies and the isolated environment the chromatin was studied in should have ensured the chromatin maintained a reliable conformation. As mentioned previously HMG-14/-17 is able to bind to nucleosomes and it is possible that in the conditions I used for my experiments these proteins remained bound to the nucleosomes whereas for previous experiments HMG-14/-17 may have been lost during the chromatin isolation or purification.

It is difficult to identify why differentiated cell chromatin should be more sensitive to nucleases than stem cell chromatin. Obvious reasons would include alterations in the levels of linker histone subtype complement, the presence of different linker histone subtypes or the acetylation of the core histones. All of these factors would be expected to alter the stability of the fibre which is a feature which may not be detected by sedimentation analysis, but may be identified by nuclease digestion. However, as far as I know no substantial differences in the levels of linker histones have been identified between differentiated and undifferentiated cells, and the level of histone acetylation has not been shown to alter during the differentiation of HL60 cells (O'Neill and Turner, 1995). The levels of HMG-14/-17 in cells is relatively low with only enough protein being present to bind to 1% of nucleosomes. From previous data suggesting that the binding of HMG-14/-17 may alter the stability of the chromatin fibre it is possible that only a small change in either the levels or distribution of these HMG proteins would be necessary to alter the nuclease sensitivity of these cells.

ES cells are highly dependent on expression from the Oct-4 gene to maintain their pluripotential phenotype. Therefore, the importance of this gene may necessitate it having an altered chromatin conformation. Using a hydrodynamic approach I was unable to show there was any alteration in the chromatin structure around this gene in agreement with our previous data.

These results suggest that although there is no obvious difference in the conformation of higher order chromatin fibre between differentiated and undifferentiated cells there maybe an alteration in the stability of the chromatin fibre. The presence of more discontinuities in the chromatin fibre from differentiated cells or the presence of altered types of linker histones may provide nucleation points for attack by nucleases and in turn these nicks will release the torsional and matrix constraints on the fibre allowing it to be digested more readily.

4. The conformation of centromeric heterochromatin

4.1 Introduction

The individual building blocks of the chromatin fibre have been studied in detail (Section 1.2), and the influence of this most fundamental level of chromatin organisation on gene expression is well understood (Section 1.5). In contrast, the structure of the higher order chromatin fibre which is likely to influence gene expression at a domain-wide level, regulation of cellular processes such as cell division and facilitate the nuclear organisation of chromatin remains an enigma. During interphase, the chromatin fibre probably adopts a 30nm filament-like structure (Thoma et al., 1979), which although frequently depicted as being ordered and static, probably has the capacity to be highly dynamic depending on its situation (Figure 4.1).

After mitosis most of the cellular chromatin decondenses to form euchromatin whilst a fraction of the chromatin is maintained as condensed heterochromatin. The factors responsible for maintaining the heterochromatin in a more condensed configuration are not known; this may involve the DNA sequence, associated non-histone proteins, subtypes of core and linker histones, and the modifications found on core histones and on the DNA. It is likely that active genes are contained in chromatin regions which are permissive for expression whilst inactive genes are in a transcriptionally non-permissive environment. Previous studies on the higher order chromatin fibre (Fisher and Felsenfeld, 1986; Caplan et al., 1987) have been unable to find any differences in the higher order chromatin structure around active and inactive genes, even though it is commonly thought that the core-histone acetylation found associated with active genes (Hebbes et al., 1988; Hebbes et al., 1992) would alter the local chromatin architecture. However, comprehensive *in vitro* studies have been unable to demonstrate any alteration in the chromatin fibre due to acetylation (McGhee et al., 1983a), suggesting that the role of core histone acetylation maybe to modulate the stability of the chromatin fibre to facilitate events such as transcription initiation and elongation rather than to directly alter its overall conformation (Annunziato et al., 1988; Tse et al., 1998).

To date, it has not been demonstrated if there is any alteration in the conformation of the 30nm chromatin fibre in different situations. As shown in Chapter 3, there appears to be no change in chromatin structure for bulk chromatin in both undifferentiated and differentiated cell types, and for loci which contain critically important regulatory genes (Figure 3.17). The studies undertaken previously in this laboratory (Kimura et al., 1983; Caplan et al., 1987) and in others (Fisher and Felsenfeld, 1986) have examined the conformation of active and inactive genes in chicken erythrocytes, but the chromatin structure of constitutive heterochromatin has not been analysed. The sedimentation technique presented in Chapter 3 lends itself to both an analysis of bulk chromatin, as presented, by studying chromatin fragments taken from sucrose gradients using ethidium bromide staining, and to the study of specific gene sequences (Figure 3.18) by the blotting of these gels and subsequent probing with sequence-specific probes. This hydrodynamic sedimentation approach can therefore be used to investigate whether there is a difference in the chromatin conformation of heterochromatic sequences.

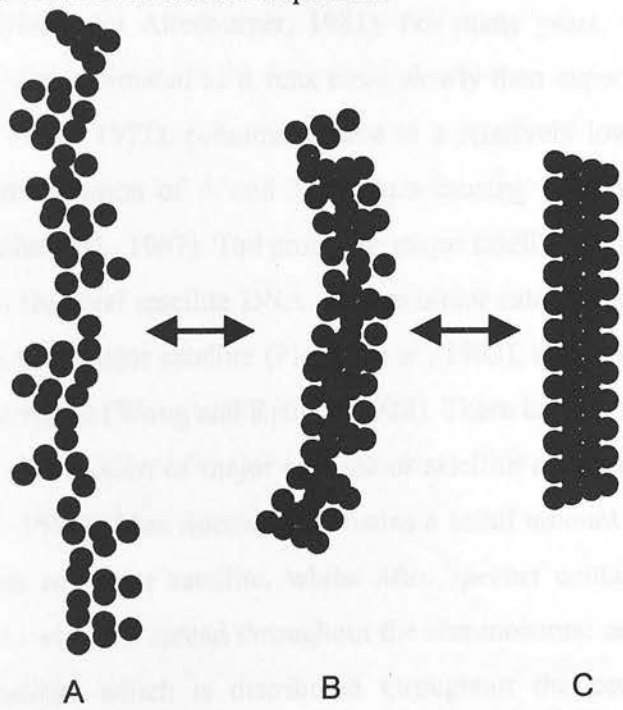


Figure 4.1 The dynamic 30nm chromatin fibre. The 30nm chromatin fibre is frequently depicted as being a rigid structure, but in reality it is probably highly dynamic and can adopt a number of conformations. A could be an active domain, B could be a poised domain and C could be an inactive region such as heterochromatin.

Of the constitutively heterochromatic regions of chromosomes (Figure 1.5), the centromere (Section 1.3) and the telomere (Section 1.4) have unique positions and roles in the formation, regulation and stability of chromosomes. All mouse centromeres, except the centromere on the Y chromosome, are acrocentric (Figure 1.6; Evans, 1981), whilst in humans the majority of chromosomes are metacentric. In higher eukaryotes the centromere regions, marked by the primary constriction, are invariably associated with an underlying array of complex satellite sequences composed of tandem repeats (Figure 1.10).

In the laboratory mouse, *Mus. musculus*, satellite sequences constitute 5-10% of the total cellular DNA (Britten and Wang 1966; Horz and Altenburger 1981), of which two centromeric satellites have been identified. The predominant centromeric satellite is the major (or γ) satellite, which was shown to be localised to the centromere region of the chromosome (Jones, 1970; Pardue and Gall, 1970). It is based on a tandem 234 base pair repeat in which almost all CpG pairs are extensively methylated (Horz and Altenburger, 1981). For many years, the size of the major satellite was over-estimated as it runs more slowly than expected in polyacrylamide gels (Zeiger et al., 1971), presumably due to a relatively low GC content, and an asymmetric distribution of A and T residues causing it to form a bent secondary structure (Radic et al., 1987). The prototype major satellite sequence is representative of 60-70% of the total satellite DNA. Mouse minor satellite is present at a level 10-20 times less than major satellite (Pietras et al., 1983), and is composed of a tandem 120 base pair repeat (Wong and Rattner, 1988). There is substantial species variation between the distribution of major and minor satellite sequences in the genus *Mus.* (Wong et al., 1990). *Mus. musculus*, contains a small amount of minor satellite and large amounts of major satellite, whilst *Mus. spretus* contains small amounts of major satellite which is spread throughout the chromosomal arm and a large amount of minor satellite which is distributed throughout the centromeric domain. In contrast, the satellite in the asian mouse, *Mus. caroli*, which diverged from *Mus. musculus* 5-7 million years ago, is not detected using minor satellite probes; it has been shown to have two abundant satellite DNA sequences of its own (Kipling et al., 1995). The minor satellite is therefore thought to have evolved from the major

satellite after the divergence of *Mus. caroli*, which then evolved its own satellite (Wong and Rattner, 1988; Wong et al., 1990). In *Mus. musculus* (subsequently referred to as mouse) the inner centromeric regions are associated with 300 kb minor satellite sequence per chromosome whilst the major satellite repeat lies distal to the minor satellite repeat (Joseph et al., 1989; Wong et al., 1990). The mouse satellite repeats are highly conserved between chromosomes indicative of frequent inter-chromosome recombination (Vissel and Choo, 1989).

The centromeres found in primates are characterised by a 171 base pair α-satellite repeat (Vissel and Choo, 1987) which has a heterogeneity of 10-40%. This repeat is substantially more diverged than the mouse centromeric repeats such that it is possible to study both the evolution of α-satellite sequences (Choo et al., 1989) and to identify centromere-specific probes.

Both mouse minor satellite, and human α-satellite contain a conserved 17 base pair binding motif for CENP-B (Masumoto et al., 1989), whilst the *Mus. caroli* satellite has an alternative CENP-B binding motif (Kipling et al., 1995). Originally, CENP-B was thought to be an important determinant for the formation of centromeres but as deletion of the protein shows no phenotypic effect (Kapoor et al., 1998; Perez-Castro et al., 1998; Kalitsis et al., 1998b) its role is currently unclear. CENP-B has been shown to dimerise (Yoda et al., 1998) so it may be involved in binding the two sister chromatids together in a semi-redundant process. The generation of a functional artificial chromosome using cloned α-satellite sequences has been shown to occur at a high frequency using a satellite which contains CENP-B boxes and a regular repeat, suggesting these features are important for centromere formation (Masumoto et al., 1998).

Although normal centromeres are associated with satellite sequences, functional neocentromeres which do not contain satellite sequences have been identified. Sequencing of a relatively small 80 kb region around a neocentromere, found on chromosome 10 (Barry et al., 1999), has shown that its associated DNA sequence is substantially different to the canonical centromere found on this chromosome

(Jackson et al., 1999). Although this neocentromere is found to be associated with a variety of centromere associated proteins including CENP-A, CENP-C and CENP-E suggesting it is functional, it does not appear to contain motifs for CENP-B, and is not associated with any identifiable DNA sequence normally found at centromeres, suggesting that the underlying sequence is unnecessary for centromere formation. This immediately questions the role of the tandem satellite repeats found at centromeres as the centromere appears to function in their absence and neocentromeres have been shown to pass through the human germline with no apparent defects (Tyler-Smith et al., 1999). It is possible that the centromere is defined by an epigenetic effect, such as the replication timing of the higher order chromatin fibre, which is maintained in the absence of centromere repeats, the role of the repeat sequences having become redundant.

At centromeres the normal complement of core histones is altered, with H3 being replaced by CENP-A to an as yet undefined extent (Sullivan et al., 1994; Shelby et al., 1997). The role of CENP-A in replacing H3 has not been identified. It has been suggested that this may change the conformation of the chromatin fibre around centromeres, either directly or by the protein having different attributes to H3, such as an alteration in CENP-A phosphorylation (Wei et al., 1999) or acetylation patterns(Chadee et al., 1999).

Centromeres are critical for the formation and stability of a functional chromosome. Understanding the molecular components required for their formation will allow us to further understand how neocentromeres and dicentrics are able to function and will facilitate the development of artificial chromosomes. To explore the functional components of centromeres, the role of satellite sequences and the identification of chromatin with an altered conformation I have examined the underlying chromatin structure of centromeric heterochromatin to determine whether it, like active genes, has a similar conformation to bulk chromatin, or whether it has a special higher order structure.

4.2 Analysis of mouse centromeric heterochromatin by micrococcal nuclease digestion

Mouse centromeric heterochromatin is characterised by two classes of satellite sequence: minor satellite located proximal to the centromere and major satellite located distal to the minor satellite (Joseph et al., 1989; Wong et al., 1990). To ensure that a representative proportion of satellite chromatin was released from micrococcal nuclease-digested nuclei, NIH3T3 nuclei were prepared and digested briefly with micrococcal nuclease to yield long polynucleosomes. Prior to releasing the chromatin from the nuclei, an aliquot (S_0) of total chromatin was removed and the soluble chromatin (S_2) released from the nuclei overnight. The DNA from the two samples was purified and equal amounts were size fractionated on an agarose gel (Figure 4.2A). The gel was Southern blotted and hybridised to both a mouse minor satellite probe (Plasmid #47; Kipling et al., 1994; Figure 4.2B), or a major satellite probe (Plasmid #77; Lewis et al., 1992; Figure 4.2C). The blots show that a representative fraction of the minor and major satellite are released from the nuclei, although release of the material is not 100% as expected from my earlier results (Table 3.1). Previous experiments have shown that satellite chromatin is preferentially precipitated in the presence of low salt and calcium ions (Horvath and Horz, 1981; Zhang and Horz, 1982). The approach used here, which releases the chromatin in TEP_{20} and in the absence of calcium, does not appear to precipitate the satellite chromatin to any significant extent.

To analyse the chromatin structure of mouse centromeric heterochromatin, soluble chromatin was prepared from NIH3T3 cells and sedimented in a 6-40% sucrose gradient in TEP_{80}. The chromatin fractions were removed from the gradient by upward displacement and were purified and size fractionated on a 1% agarose gel in TBE (Figure 4.3A). The gel was Southern blotted onto a nylon membrane and probed with a minor satellite probe (Figure 4.3B), a major satellite probe (Figure 4.3C), and a B2 interspersed repeat probe (Plasmid #84; Chambers et al., 1997) which hybridises to bulk chromatin (Figure 4.3D). The gel was analysed on a laser scanner at 473nm with a 580nm band-pass filter and the blots were analysed on a

phosphorimager. The size of the DNA peak was determined for each fraction using analysis software which had been calibrated with end-labelled DNA size markers on the gel. The size of the DNA for each fraction was plotted against the fraction numbers which, as discussed in Chapter 3, is an indication of the relative sedimentation value for the samples. By plotting the data from all of the gels it is possible to determine whether the satellite chromatin sediments at a different rate to the bulk chromatin (Figure 4.3E). The results show that the profile for the ethidium bromide stained gel lies directly on the profile for the B2 probed blot, indicating that ethidium bromide staining is representative of the sedimentation of the bulk chromatin. In contrast, the profiles for the sedimentation of the major and minor satellite lie below this line. These results show that for two equally sized DNA fragments (reading across the graph), minor satellite and major satellite sediment faster than bulk in the sucrose gradient. The simplest interpretation of these results is that centromeric chromatin has a more compact higher order chromatin structure than bulk chromatin. In addition, these results could suggest that this degree of compaction decreases away from the centromere as the sedimentation of the more distal major satellite lies between the bulk and proximally located minor satellite. This experiment was repeated on four different chromatin preparations and found to be consistent. There was some variation from preparation to preparation depending on the size of the chromatin fragments, as larger chromatin fragments were found to further accentuate these differences.

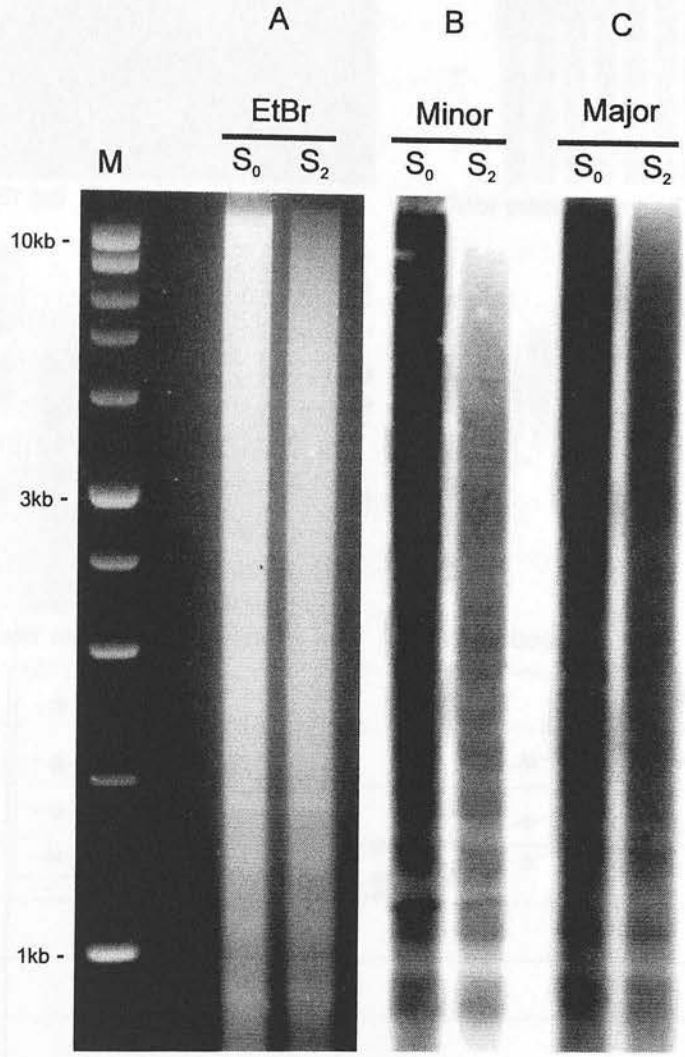


Figure 4.2 Representative release of satellite chromatin from NIH3T3 nuclei. An aliquot of DNA was purified from total NIH3T3 nuclei (S_0) and an aliquot of DNA was purified from soluble released NIH3T3 chromatin (S_2). The DNA was fractionated on a 1% agarose gel. The gel was Southern blotted and probed for minor satellite and major satellite.

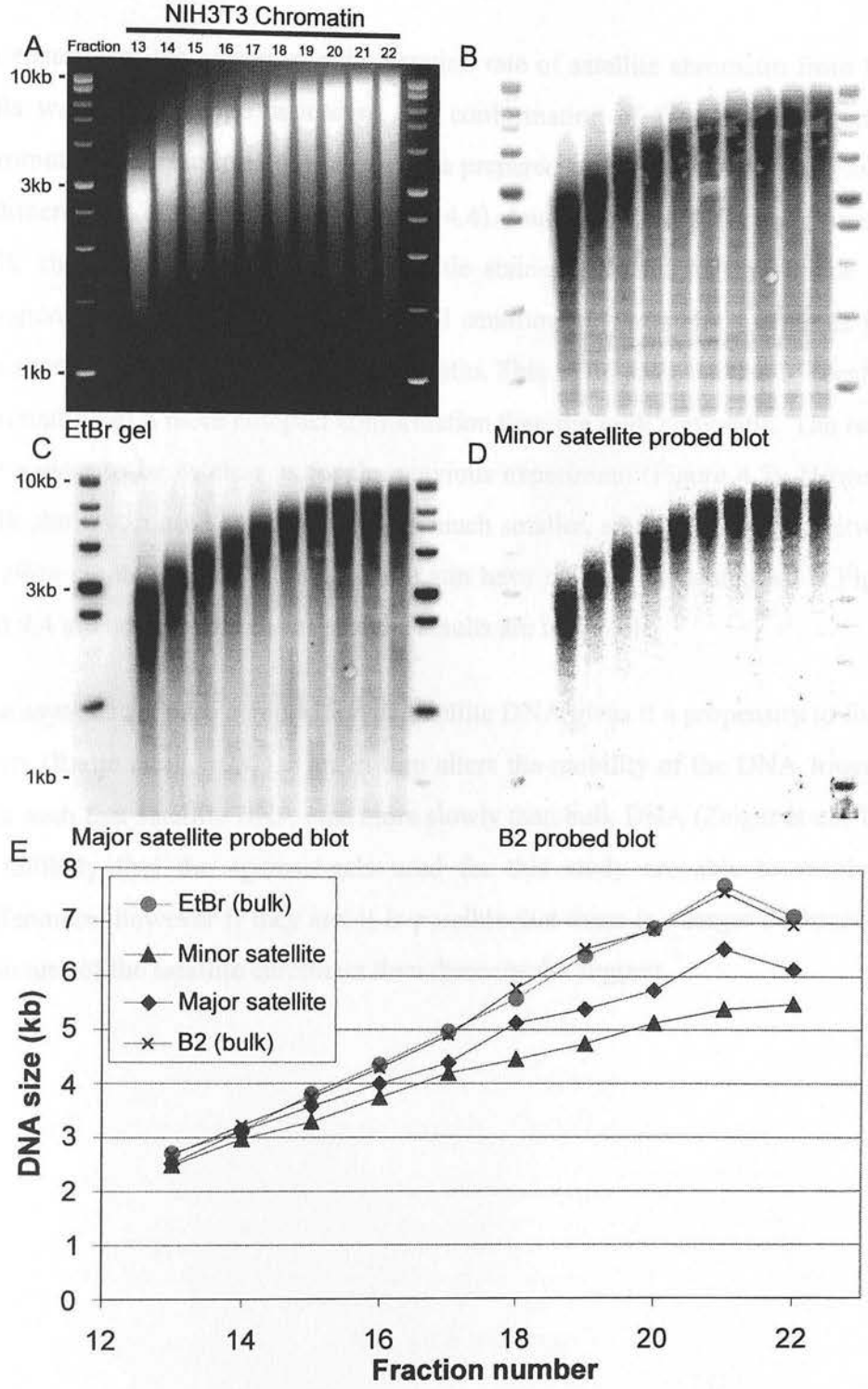


Figure 4.3 NIH3T3 centromeric heterochromatin is more compact than bulk chromatin and this compaction decreases away from the centromere. (A) Soluble chromatin from NIH3T3 cells was analysed by the sedimentation/blotting approach (see text for details) and probed for minor satellite (B), major satellite (C) and with a B2 repeat (D). (E) Comparison between the DNA size and sedimentation for bulk and satellite chromatin indicates that for two chromatin fragments of the same size the satellite chromatin sediments faster than the bulk. Markers are 1 kb ladder.

To ensure that the increased sedimentation rate of satellite chromatin from NIH3T3 cells was not cell type dependent, the conformation of F9 stem cell centromeric chromatin was examined. Chromatin was prepared from F9 cells and analysed by the sedimentation/blotting approach (Figure 4.4). Analysis of the DNA peak sizes for the bulk chromatin in the ethidium bromide stained gel and for the minor satellite component in the blot show that F9 cell centromeric chromatin sediments faster in the sucrose gradient that the bulk chromatin. This again suggests that F9 centromeric chromatin has a more compact conformation than the bulk chromatin. The results do not appear to be as clear as for the previous experiment (Figure 4.3). However, the bulk chromatin used in this analysis is much smaller, so the differences between the satellite chromatin and the bulk chromatin have not been accentuated. If Figure 4.3 and 4.4 are compared size-for-size the results are identical.

The asymmetric base composition of satellite DNA gives it a propensity to form bent DNA (Radic et al., 1987). This in turn alters the mobility of the DNA fragments in gels such that satellite DNA runs more slowly than bulk DNA (Zeiger et al., 1971). It is unlikely that the agarose gels used for this study are able to resolve these differences, however if they are, it is possible that there is a larger difference in the structure of the satellite chromatin than these results suggest.

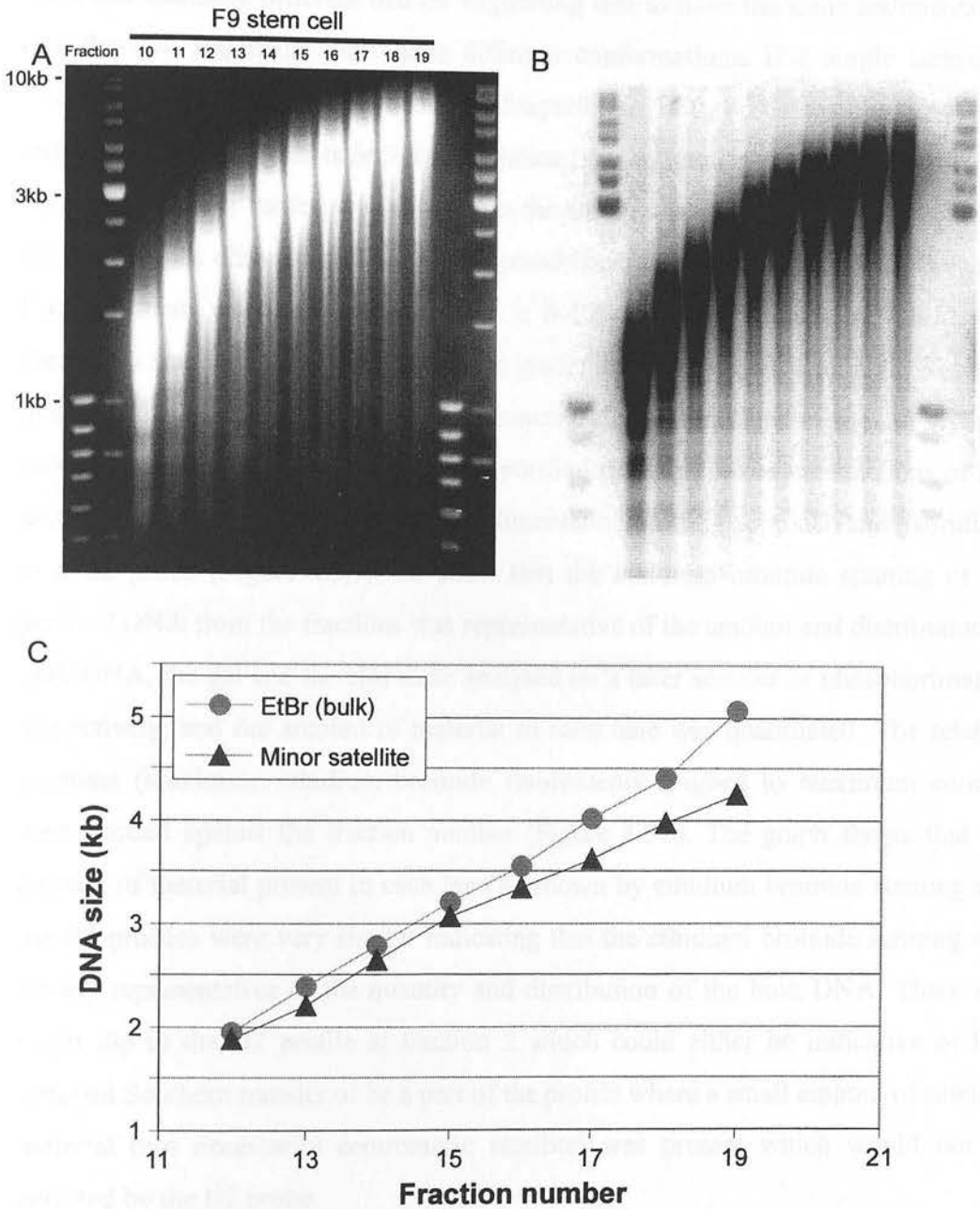


Figure 4.4 F9 stem cell centromeric heterochromatin is more compact than bulk chromatin. (A) F9 chromatin was analysed by the sedimentation/blotting approach and hybridised to a minor satellite probe (B). Comparison between the DNA size and the extent of sedimentation for bulk and satellite chromatin indicates that for two chromatin fragments of the same size the satellite chromatin sediments faster than the bulk chromatin. Markers are 1 kb ladder and 100 bp ladder.

Another way to consider these results is that for two chromatin fragments, one bulk and one satellite, in a single isokinetic fraction, the two fragments will have different sizes and therefore different masses suggesting that to have the same sedimentation rate the two fragments must have different conformations. If a single isokinetic fraction is isolated from a gradient and dialysed into TEP_5 buffer the chromatin will unfold, losing its higher order characteristics. The subsequent sedimentation of this fragment in TEP_5 buffer should separate the smaller satellite sequences away from the larger bulk chromatin. To test this possibility, soluble chromatin was prepared from F9 stem cells and sedimented in a 6-40% sucrose gradient in TEP_{80}. The chromatin fractions were pumped off the gradient and individual isokinetic fractions dialysed into TEP_5. The chromatin was concentrated and then sedimented in separate 6-40% sucrose gradients in TEP_5. DNA purified from the chromatin fractions of this second gradient were analysed by the sedimentation/blotting approach and hybridised to a B2 probe (Figure 4.5B). To show that the ethidium bromide staining of the purified DNA from the fractions was representative of the amount and distribution of bulk DNA, the gel and the blot were analysed on a laser scanner or phosphorimager, respectively, and the amount of material in each lane was quantitated. The relative amounts (maximum ethidium bromide fluorescence aligned to maximum counts) were plotted against the fraction number (Figure 4.5C). The graph shows that the amount of material present in each lane as shown by ethidium bromide staining and the B2 profiles were very similar indicating that the ethidium bromide staining was closely representative of the quantity and distribution of the bulk DNA. There is a slight dip in the B2 profile at fraction 2 which could either be indicative of less efficient Southern transfer or be a part of the profile where a small amount of satellite material (not necessarily centromeric satellite) was present which would not be detected by the B2 probe.

The fractions from another TEP_5 gradient were taken and analysed on a 1% agarose gel (Figure 4.5D). The gel was Southern blotted and this time hybridised to a minor satellite probe (Figure 4.5E). Once again the amount of bulk DNA present in each lane of the ethidium bromide stained gel, and the amount of minor satellite DNA present in each lane of the blot, were quantitated and plotted against each other

(Figure 4.5F). The graph shows that the minor satellite chromatin is distinctly separated away from the bulk chromatin in the TEP_5 gradient and as it sediments more slowly it must be smaller than the bulk chromatin. This demonstrates that in an individual isokinetic fraction from the original sucrose gradient run in TEP_{80}, the centromeric chromatin is smaller than the bulk chromatin. Thus, in the TEP_{80} gradient to have the same sedimentation rate as the bulk chromatin, the satellite chromatin must have a more compact conformation.

This result also demonstrates that the increased rate of sedimentation observed for the satellite chromatin is not due to aggregation of the satellite material in the TEP_{80} sucrose gradient because when the fibre is unfolded, by dialysis into low salt, the bulk and satellite chromatins clearly have different sizes. There is still a slight difference in the DNA sizes for the bulk chromatin (Figure 4.5D) and the satellite chromatin (Figure 4.5E) in individual equivalent fractions (e.g. fraction 10). This is probably indicative of the chromatin fibre not being completely unfolded in TEP_5 (Thoma et al., 1979). This experiment was repeated with another fraction and the same blot was probed with minor satellite, major satellite and B2 giving the same results (data not shown).

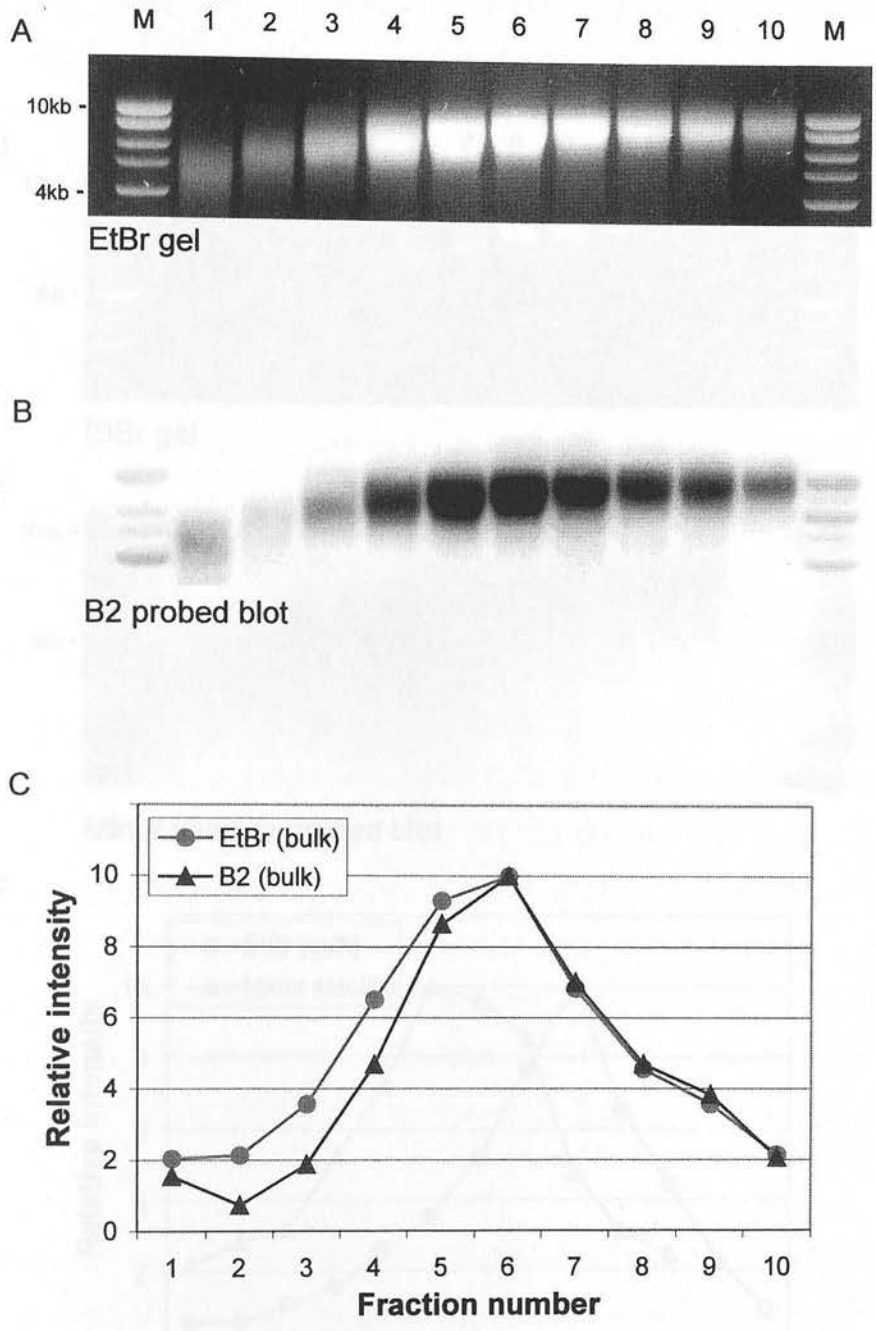
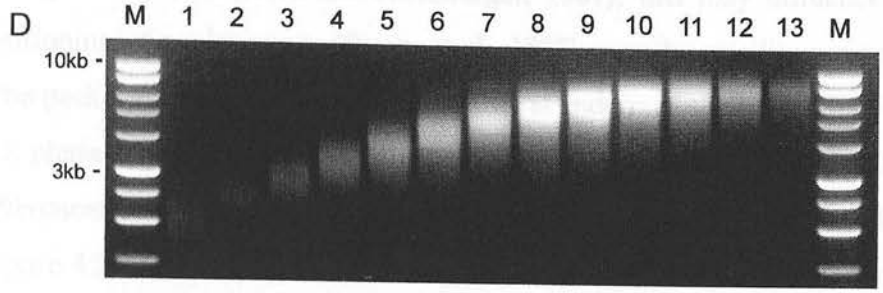
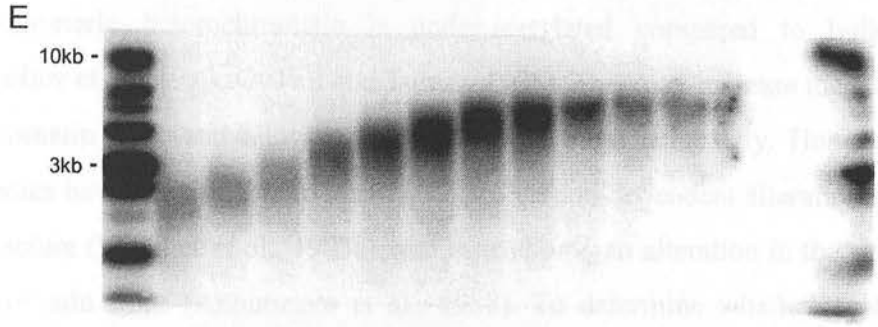


Figure 4.5 In an isokinetic TEP_{80} fraction, F9 stem cell satellite chromatin is more compact than bulk chromatin. F9 stem cell chromatin was fractionated in a folded conformation in a sucrose gradient containing TEP_{80}. Individual isokinetic fractions were dialysed into TEP_5 and fractionated by size on a sucrose gradient in TEP_5. (A) Aliquots from the TEP_5 gradient were purified and fractionated on an agarose gel. (B) The gel was Southern blotted with a B2 probe. (C) Analysis of the gel (A) and the blot (B) shows that ethidium bromide staining is equivalent to bulk DNA (D). (cont.)



EtBr gel



Minor satellite probed blot

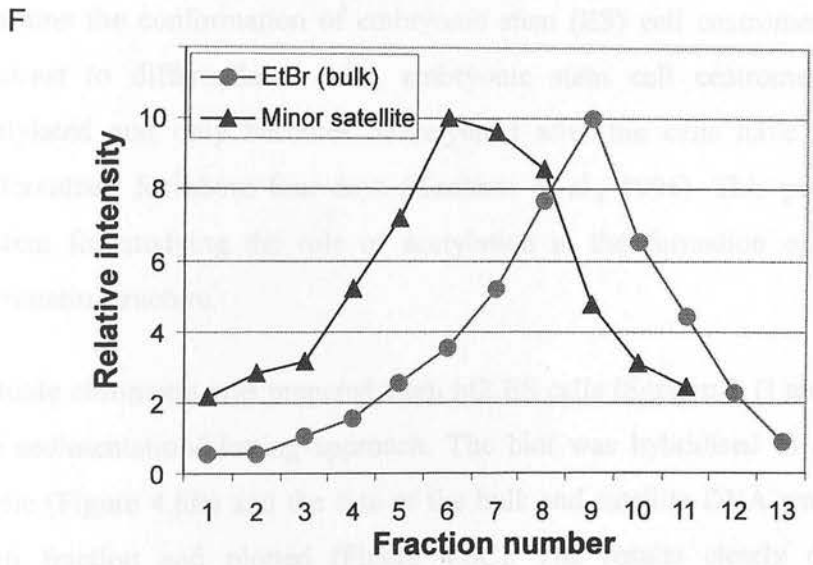


Figure 4.5 (cont.) (E) Another fraction was analysed in a similar manner but hybridised to a minor satellite probe. (F) Analysis of the gel and blot shows that the satellite chromatin is smaller and therefore more compact than bulk chromatin. Markers are 1 kb ladder.

Possible reasons for the altered compaction of satellite chromatin include modification of the underlying DNA or histone proteins. Satellite chromatin is heavily methylated (Horz and Altenburger, 1981); this may influence the regular positioning of nucleosomes (Davey et al., 1997) over the satellite region allowing it to be packaged more compactly. F9 satellite is undermethylated and replicates earlier in S phase (Selig et al., 1988) than the satellite found in NIH3T3 cells. Despite these differences both cell types appear to have similarly compacted centromeric chromatin (Figure 4.3 and 4.4) suggesting that DNA methylation is not entirely responsible for the altered chromatin structure found at centromeres. Another possibility is that centromeric heterochromatin is under-acetylated compared to bulk chromatin (Pashev et al., 1983; O'Neill and Turner, 1995). This may increase the stability of the chromatin fibre and allow it to be packaged more compactly. However, previous studies have been unable to identify an acetylation-dependent alteration in chromatin structure (McGhee et al., 1983a), and instead only an alteration in the stability of the chromatin fibre (Annunziato et al., 1988). To determine whether acetylation was responsible for the more compact centromeric chromatin structure, I decided to examine the conformation of embryonic stem (ES) cell centromeric chromatin. In contrast to differentiated cells, embryonic stem cell centromeric chromatin is acetylated and only becomes deacetylated after the cells have been induced to differentiate for about four days (Keohane et al., 1996). This provides a suitable system for studying the role of acetylation in the formation of the centromeric chromatin structure.

Soluble chromatin was prepared from ht2 ES cells (Section 3.1) and analysed using the sedimentation/blotting approach. The blot was hybridised to a minor satellite probe (Figure 4.6B) and the size of the bulk and satellite DNA was determined for each fraction and plotted (Figure 4.6C). The results clearly demonstrate that equivalently sized embryonic stem cell minor satellite chromatin fragments sediment faster than bulk chromatin fragments showing that they too have a more compact higher order chromatin conformation. Overall, these results suggest that the altered chromatin conformation observed at centromeric sequences is unlikely to be a direct consequence of differences in either histone acetylation or DNA methylation.

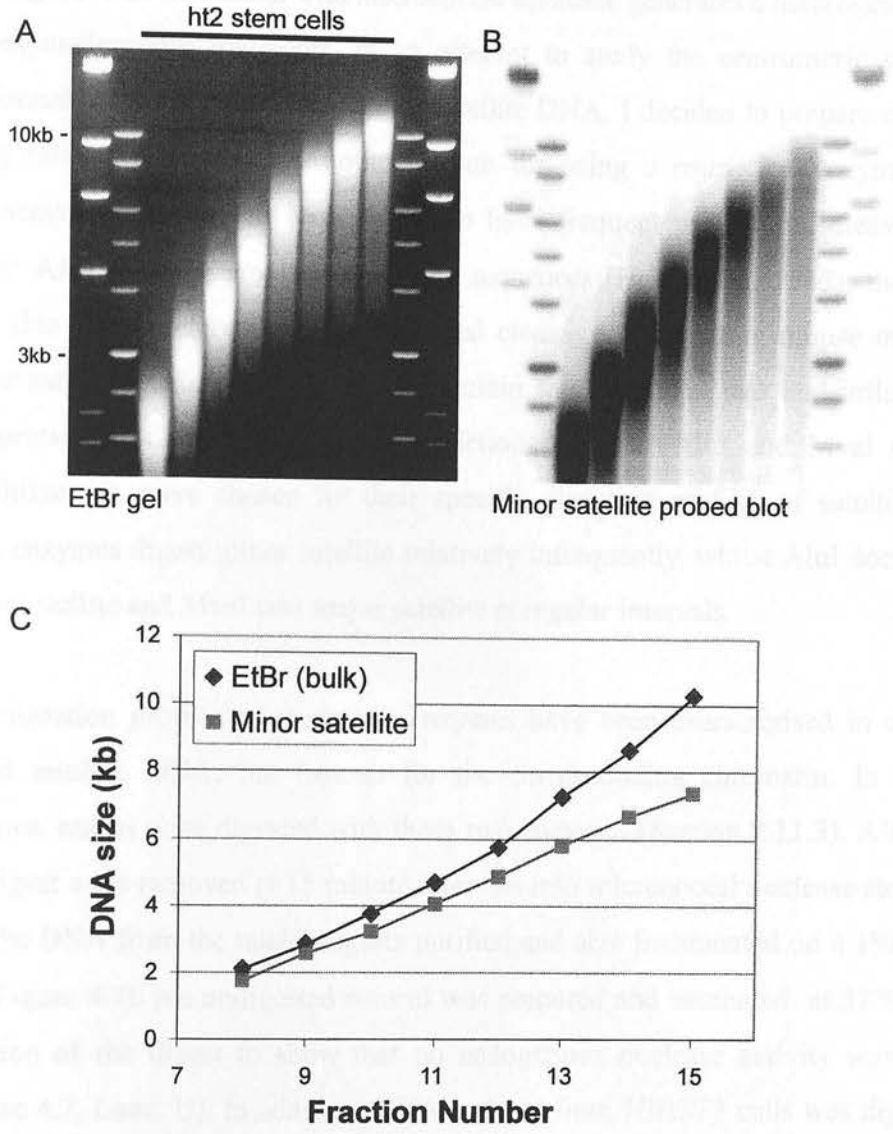


Figure 4.6 Centromeric heterochromatin compaction appears to be independent of histone acetylation. (A, B) ht2 stem cell chromatin was analysed by the sedimentation/blotting approach and hybridised to a minor satellite probe. (C) Equivalently sized ES cell satellite chromatin sediments faster than bulk chromatin suggesting that the acetylated satellite chromatin is more compact than bulk chromatin. Markers are 1 kb ladder and λ HindIII

4.3 Analysis of mouse centromeric heterochromatin by restriction enzyme digestion

The digestion of chromatin with micrococcal nuclease generates a heterogeneous mix of polynucleosome fragments. In an attempt to study the centromeric chromatin conformation of discrete fragments of satellite DNA, I decided to prepare chromatin using restriction enzymes. Another reason for using a restriction enzyme is that micrococcal nuclease has been shown to have frequent preferential cleavage sites within African green monkey α-satellite sequences (Horz et al., 1983); therefore it may also have preferential and undefined cleavage sites within mouse major and minor satellite which could select for certain satellite sequences and influence the interpretation of the results. Two restriction enzymes, AluI and MvaI (a BstNI isoschizomer), were chosen for their specific digestion profiles of satellite DNA. Both enzymes digest minor satellite relatively infrequently, whilst AluI does not cut major satellite and MvaI cuts major satellite at regular intervals.

The digestion profiles of these two enzymes have been characterised in detail for naked satellite DNA, but less so for the corresponding chromatin. In the first instance, nuclei were digested with these two enzymes (Section 2.11.3). Aliquots of the digest were removed at 15 minute intervals into micrococcal nuclease stop buffer and the DNA from the nuclei digests purified and size fractionated on a 1% agarose gel (Figure 4.7). An undigested control was prepared and incubated at 37°C for the duration of the digest to show that no endogenous nuclease activity was present (Figure 4.7, Lane: U). In addition, DNA isolated from NIH3T3 cells was digested to completion using AluI or MvaI (Figure 4.7, Lane: Naked AluI, Naked MvaI). In the MvaI digest it is possible to see a monomer, dimer and trimer etc. repeat of the major satellite. In the AluI digest there are no distinct satellite bands as expected. A nucleosomal repeat is evident in the AluI digest of the chromatin, whilst in the MvaI digest the nucleosomal repeat can be seen to be overlaid with the major satellite repeat (a similar gel was Southern blotted and probed with a major satellite probe to confirm these bands were major satellite (data not shown)). As has been previously shown for African green monkey (Zhang et al., 1983) and mouse (Zhang and Horz,

1984) the nucleosomal repeats do not lie directly in phase with the satellite repeats but the nucleosomes do occupy a number of discretely defined positions. In mouse major satellite, the sixteen preferred positions correlate with an internal highly diverged 9 base-pair subrepeat (Zhang and Horz, 1984) of the satellite DNA. This suggests that the satellite repeat may provide a template for the regular positioning of nucleosomes allowing this sequence to be packaged in a regular and therefore more compact manner.

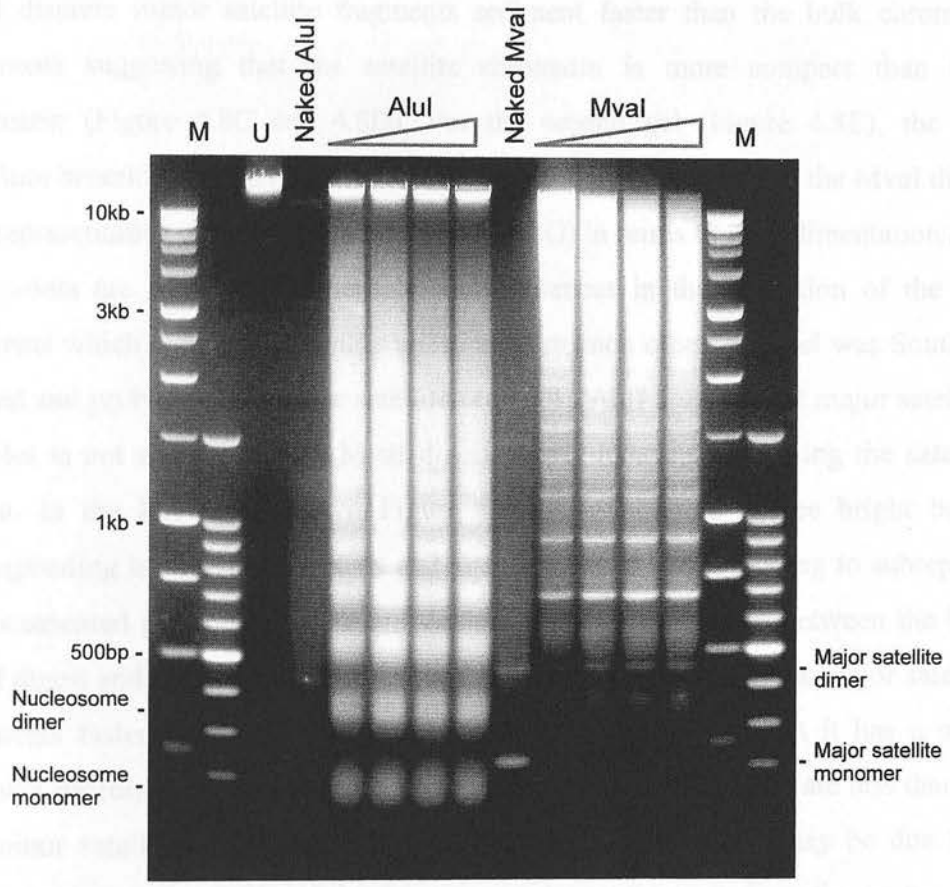


Figure 4.7 Chromatin digestion using restriction enzymes reveals the major satellite repeat pattern. NIH3T3 nuclei were digested with either AluI or MvaI. The DNA was purified and analysed by agarose gel electrophoresis. Markers are 1 kb and 100 bp ladder.

To analyse the chromatin conformation of discrete satellite sequences, soluble chromatin was prepared from NIH3T3 nuclei using MvaI or AluI (Section 2.7.4) and the soluble chromatin was released overnight. The chromatin was sedimented in a 6-40% sucrose gradient in TEP_{80} and was isolated from the gradients and fractionated on 1% agarose gels in TBE (Figure 4.8A and 4.8E). The gels were Southern blotted and one of the gels was probed with a minor satellite probe (Figure 4.8B). Analysis of these results is slightly less precise than the previous analysis as the lower fractions are more difficult to size as they are a mix of discrete bands. Analysis of both the AluI and MvaI digests probed with minor satellite show that equivalently sized discrete minor satellite fragments sediment faster than the bulk chromatin fragments suggesting that the satellite chromatin is more compact than bulk chromatin (Figure 4.8C and 4.8D). For the second gel (Figure 4.8E), the two ethidium bromide stained digests were compared to demonstrate that the MvaI digest was representative of bulk chromatin (Figure 4.8G) in terms of its sedimentation rate. The points are offset to compensate for differences in the collection of the two gradients which were out of synchronisation from each other. The gel was Southern blotted and probed with a major satellite probe. As AluI does not cut major satellite, the blot is not shown, but the MvaI digest clearly hybridises showing the satellite repeat. In the left-hand lane of Figure 4.8F it is possible to see bright bands corresponding to individual repeats and ½ and ¼ bands corresponding to subrepeats as documented previously (Horz and Zachau, 1977). Comparison between the bulk MvaI digest and the major satellite-probed MvaI digest shows that the major satellite sediments faster than equivalently sized bulk DNA indicating that it has a more compact conformation than the bulk chromatin. The differences seen are less than for the minor satellite probed blots (Figure 4.8C and 4.8D) which may be due to a decrease in the compaction of the centromeric chromatin away from the centromere (Figure 4.3E). These results demonstrate that the differences observed in the compaction of the centromeric chromatin can be recreated using discretely sized chromatin fragments reinforcing the observation that centromeric heterochromatin is more compact than bulk chromatin.

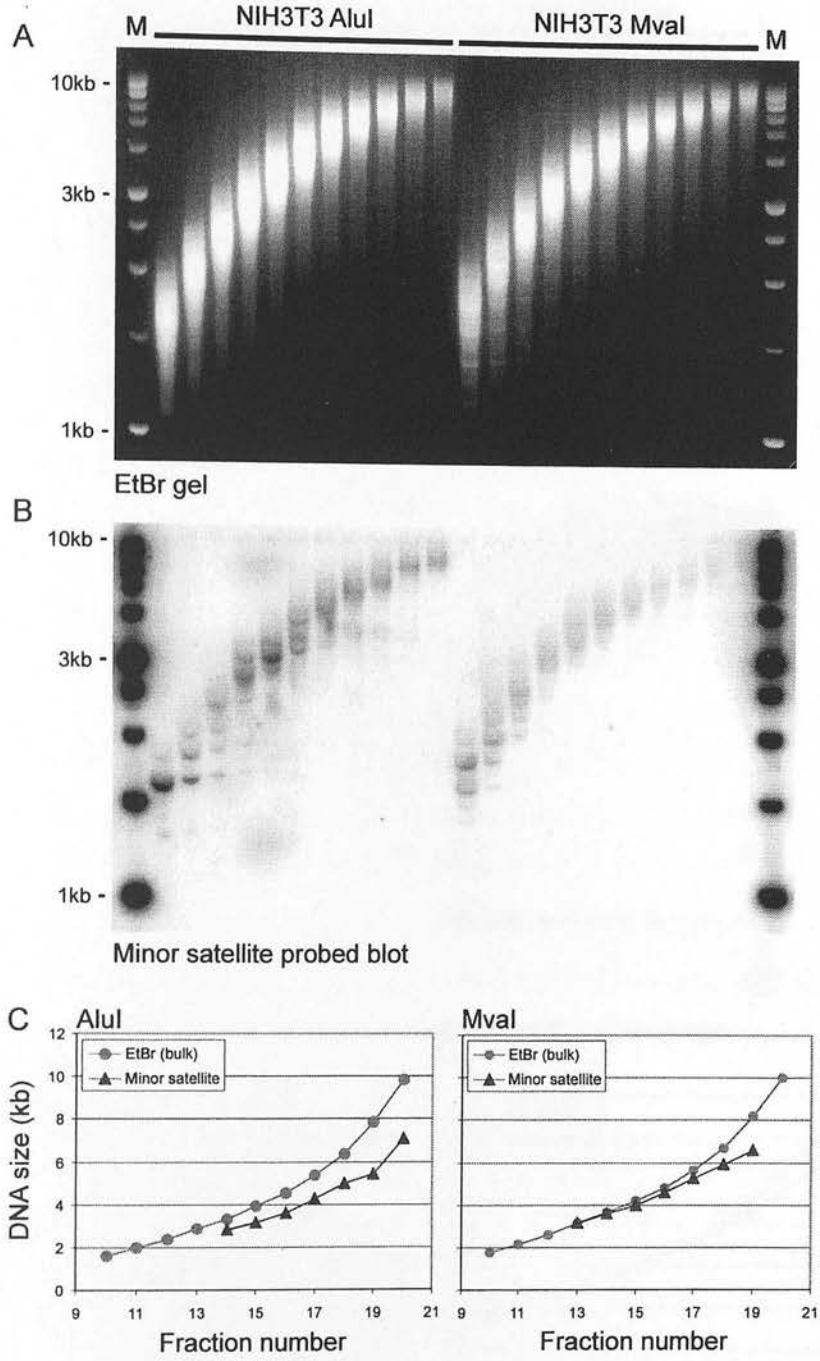


Figure 4.8 The conformation of NIH3T3 centromeric heterochromatin is more compact than bulk chromatin, as determined by the sedimentation of discrete restriction fragments. (A, D) NIH3T3 nuclei were digested with AluI or MvaI. The chromatin was analysed by sedimentation and blotting. (B) Hybridisation with a minor satellite probe reveals the minor satellite banding pattern. (C) Analysis shows that for two chromatin fragments of equal size the minor satellite chromatin sediments more rapidly. (cont.)

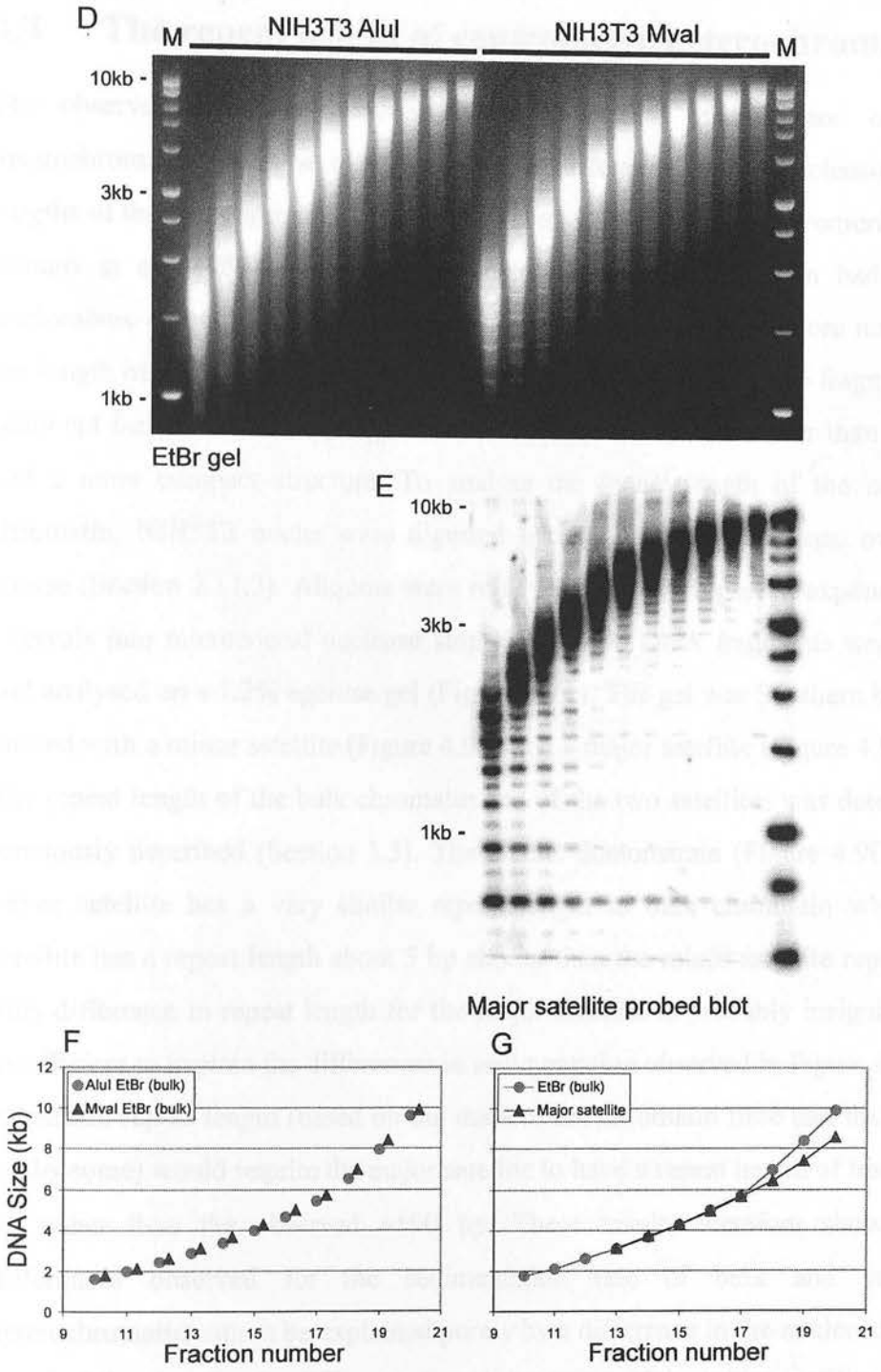


Figure 4.8 (cont.) (F) Comparisons between the bulk AluI and MvaI digests indicate that the digests release equivalently sedimenting fragments. (E) Hybridisation with a major satellite probe clearly reveals the major satellite banding pattern. Analysis (G) shows that for two chromatin fragments of equal size the major satellite chromatin sediments more rapidly than bulk chromatin suggesting it has a more compact conformation. Markers are a 1 kb ladder.

4.4 The repeat length of centromeric heterochromatin

The observed difference between the sedimentation of bulk and centromeric heterochromatin could be a reflection of a difference in the nucleosome repeat lengths of the two chromatins, as has been shown for other pericentromeric satellites (Omori et al., 1980). Essentially, if centromeric heterochromatin had a shorter nucleosome repeat length than the bulk chromatin it would have more nucleosomes per length of DNA, increasing the mass of the fragment. In turn, the fragment would sediment faster in the sucrose gradient because it was heavier rather than because it had a more compact structure. To analyse the repeat length of the centromeric chromatin, NIH3T3 nuclei were digested with micrococcal nuclease over a time course (Section 2.11.2). Aliquots were removed from the digest at exponential time intervals into micrococcal nuclease stop buffer. The DNA fragments were purified and analysed on a 1.2% agarose gel (Figure 4.9A). The gel was Southern blotted and probed with a minor satellite (Figure 4.9B) and a major satellite (Figure 4.9C) probe. The repeat length of the bulk chromatin and of the two satellites was determined as previously described (Section 3.3). The results demonstrate (Figure 4.9D) that the minor satellite has a very similar repeat length to bulk chromatin whilst major satellite has a repeat length about 5 bp shorter than the minor satellite repeat length. This difference in repeat length for the major satellite is probably insignificant and insufficient to explain the differences in sedimentation observed in Figure 4.3E as the calculated repeat length (based on the mass of the chromatin fibre and the mass of a nucleosome) would require the major satellite to have a repeat length of less than 152 bp rather than the observed ≈190 bp. These results therefore show that the differences observed for the sedimentation rate of bulk and centromeric heterochromatin cannot be explained purely by a difference in the nucleosome repeat length. Likewise, for the difference in sedimentation of the minor satellite chromatin fibre to be explained by non-nucleosomal protein mass, the satellite fibre would have to be 34% heavier by protein mass which is unlikely, if not unfeasible. It has also been shown that mouse satellite DNAs and bulk DNAs have the same protein to DNA ratio by mass (Mazrimas et al., 1979).

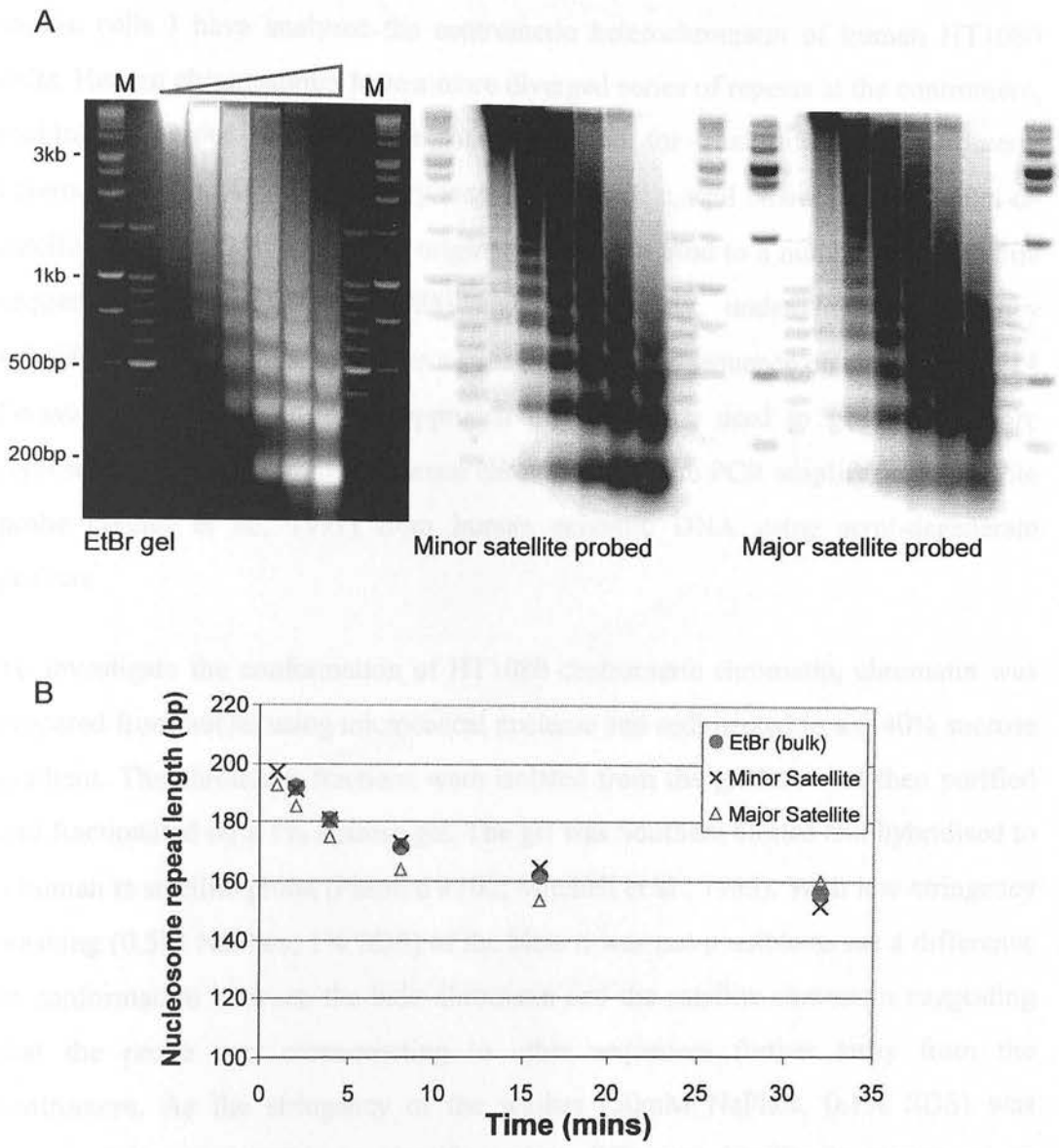


Figure 4.9 Centromeric heterochromatin has a similar repeat length to bulk chromatin. (A) NIH3T3 nuclei were digested with micrococcal nuclease. The DNA was purified and size fractionated. The gel was blotted and hybridised to a minor satellite or major satellite probe. (B) The repeat lengths of the bulk chromatin and the satellite chromatins were determined. Markers are 1 kb and 100 bp ladders.

4.5 Analysis of human centromeric heterochromatin

My observations suggest that *Mus. musculus* centromeric heterochromatin is more compact than bulk chromatin from the same cell type. To ensure this is not unique to mouse cells I have analysed the centromeric heterochromatin of human HT1080 cells. Human chromosomes have a more diverged series of repeats at the centromere, making it harder to identify a suitable probe for analysing the centromeric heterochromatin. As a preliminary step, I have used a well characterised human α-satellite probe, p82H, which was originally shown to bind to a number of α-satellite sequences (Mitchell et al., 1985) and subsequently, under higher stringency hybridisation conditions, to bind to a specific α-satellite sequence on chromosome 14 (Waye et al., 1988). Another approach that could be used to generate a more representative probe for all the human chromosomes is to PCR amplify an α-satellite probe (Weier et al., 1991) from human genomic DNA using semi-degenerate primers.

To investigate the conformation of HT1080 centromeric chromatin, chromatin was prepared from nuclei using micrococcal nuclease and sedimented in a 6-40% sucrose gradient. The chromatin fractions were isolated from the gradient and then purified and fractionated on a 1% agarose gel. The gel was Southern blotted and hybridised to a human α-satellite probe (Plasmid #106; Mitchell et al., 1985). With low stringency washing (0.5M NaPhos, 1% SDS) of the blots it was not possible to see a difference in conformation between the bulk chromatin and the satellite chromatin suggesting that the probe was cross-reacting to other sequences further away from the centromere. As the stringency of the washes (50mM NaPhos, 0.1% SDS) was increased it was possible to identify a clear difference (similar to mouse minor satellite) between the sedimentation of the α-satellite and the bulk chromatin (Figure 4.10). However, I was unable to determine whether the probe was hybridising to a representative fraction of the α-satellite or if it was hybridising to its cognate sequence on chromosome 14. The p82H probe has been shown to work well for fluorescence in situ hybridisations, but it is quite possible that it may not be the most suitable probe for analysing the chromatin conformation of centromeric

heterochromatin and in future studies it will be necessary to try other α-satellite probes. Though, for these preliminary experiments, it is clear that there is a difference between the sedimentation rate of α-satellite chromatin and bulk chromatin which may, in humans, be more variable between chromosomes.

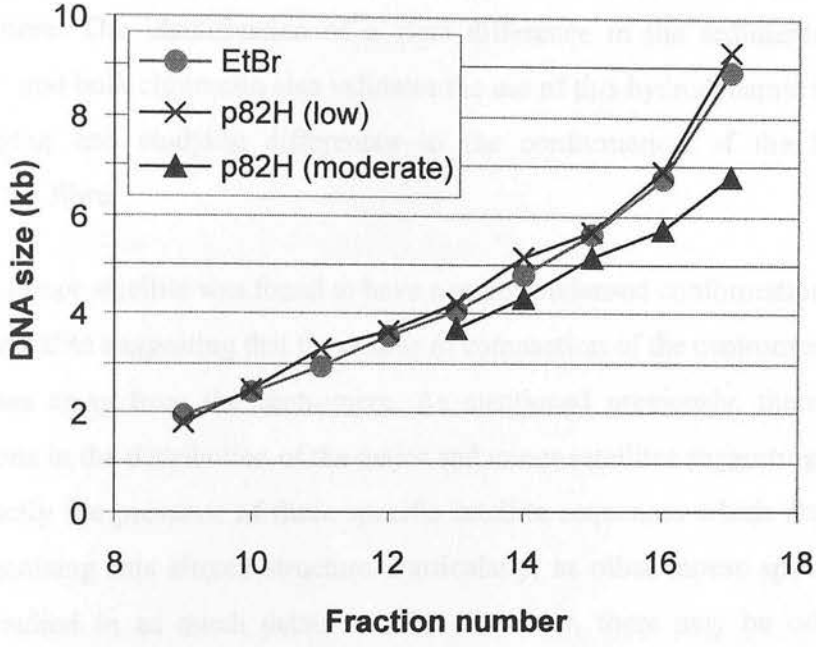


Figure 4.10 Human centromeric chromatin is more compact than bulk chromatin. HT1080 chromatin was analysed by the sedimentation/blotting approach with an α-satellite repeat. The blot was washed at low and moderate stringency (see text for details). With moderate stringency washing the satellite probe identified chromatin fragments which sedimented more rapidly than bulk chromatin suggesting they had a more compact conformation.

4.6 Discussion

The satellite chromatin associated with centromeres appears to sediment faster in sucrose gradients than equivalently sized bulk chromatin fragments, suggesting that the centromeric higher order chromatin fibre is packaged in a more compact manner. Previous studies of active and inactive genes, for example globin and ovalbumin, and of differentiated and undifferentiated cells (Chapter 3) have been unable to demonstrate any difference between the packaging of higher order chromatin fibres. Therefore, these are the first data to demonstrate an alteration in the conformation of the higher order chromatin fibre and this degree of centromeric chromatin compaction is likely to be important for the regulation and maintenance of a stable centromere. The identification of a clear difference in the sedimentation rate of satellite and bulk chromatin also validates the use of this hydrodynamic technique for identifying and studying differences in the conformation of the higher order chromatin fibre.

Mouse minor satellite was found to have a more condensed conformation than mouse major satellite suggesting that the degree of compaction of the centromeric chromatin decreases away from the centromere. As mentioned previously, there are species variations in the distribution of the major and minor satellites suggesting that it is not necessarily the presence of these specific satellite sequences which are responsible for organising this altered structure. Particularly, as other mouse species have not been studied in as much detail as *Mus. musculus*, there may be other diverged satellite sequences present as for *Mus. caroli* (Kipling et al., 1995) which are also associated with condensed centromeric heterochromatin. One of the next steps in furthering our understanding of the compaction of centromeric heterochromatin will be to investigate the centromeric chromatin conformation of related mouse species to explore the evolutionary conservation of a compact centromeric chromatin structure.

Many previous studies have isolated satellite chromatins to a high degree of purity (70-100%) using restriction enzymes and/or column-based purification approaches from a variety of species including calf (Lipchitz and Axel, 1976), rat (Igo-Kemenes et al., 1977), African green monkey (Musich et al., 1977), mouse (Zhang and Horz,

1982) and specifically from centromeric heterochromatin (Lica and Hamkalo, 1983). This has allowed the protein composition of satellite chromatin to be analysed (Mazrimas et al., 1979; Mathew et al., 1981) and characterised (Pashev et al., 1983). Recently, monkey α-satellite has been isolated using nucleoprotein hybridisation and affinity purification giving a high level of purity suitable for protein characterisation (Jasinskas and Hamkalo, 1999). My work extends previous approaches and demonstrates that it is also possible to separate centromeric satellite chromatin away from bulk chromatin based on its higher order conformation (Figure 4.5).

This result provides a means to isolate chromatin fragments which have both more and less folded conformations compared to bulk DNA and would allow the cloning and sequence evaluation of chromatin fragments with an altered higher order structure. In addition, the chromosomal location of fragments could be determined to explore changes in the chromatin structure across chromosomes.

Two further observations are apparent from these experiments. Firstly, in many of the ethidium bromide stained gels and B2 probed blots, the bulk chromatin runs as a tight band whilst the satellite-probed blots are more smeary. This is clearly seen for the major satellite probed blot where the chromatin was obtained by MvaI digestion (Figure 4.8F). If the bulk chromatin is nicked during nuclease digestion (both for micrococcal nuclease and restriction digestion), it is more likely to fall apart if the chromatin fibre has a more open conformation (Figure 4.1A). Therefore, during the sedimentation step the two fragments will sediment independently at their correct sizes. Conversely, if the satellite is more tightly packaged (Figure 4.1C) the chromatin fibre is less likely to fall apart during the sedimentation step and the fragments will sediment at apparently incorrect positions. However, after DNA purification the two DNA fragments will come apart resulting in a more smeary band than for the bulk chromatin. This is clearly seen in Figure 4.8F where there are low molecular weight satellite bands throughout the gradient when it would be impossible for them to sediment in these positions. This is presumably because a monomer, dimer or trimer etc. is cut off the larger fragment but remains attached and only becomes detached during the purification of the DNA fragment. If this is the

the chromatin conformation of human chromosome-specific satellites, which would not be possible in mouse where the satellites are highly conserved. This will necessitate isolating individual human chromosomes in human/hamster hybrid cells and undertaking an examination of the conformation of the chromatin using very specific probes. African green monkey has a similar centromeric α-satellite sequence to humans and may provide a more suitable system for analysing the chromatin conformation of a number of centromeres depending on the complexity of the underlying satellite DNA. The only disadvantage with using such a system is that the centromeric satellite arrays are far less well characterised than in humans and fewer probes are available.

Having identified a difference in the conformation of the chromatin fibre at centromeres it is necessary to determine the molecular basis for these results. The formation of a compact chromatin structure is likely to result from a combination of DNA sequence and a subset of histone or non-histone proteins. Centromeric satellite sequences have been shown to position nucleosomes precisely *in vivo* (Zhang et al., 1983; Zhang and Horz, 1984) and *in vitro* (Linxweiler and Horz, 1985) and this may allow an ordered, almost crystalline, packaging of the chromatin fibre. It is not the presence of a specific type of satellite sequence (e.g. major or minor) that generates this structure, but there is likely to be a diverged class of sequences, as found between species, which would all be suitable for the precise positioning of nucleosomes. This more condensed structure may act as a nucleating point (possibly by its late replication) for the assembly of histone and non-histone proteins which are required to maintain this compact structure. One of the principal candidates for generating an altered structure at centromeres is CENP-A. It is found at all active centromeres where it partially replaces histone H3. Alternatively, there may be a linker histone variant which is present at centromeres and maintains the compact conformation of the satellite chromatin. To identify the protein components responsible for this altered chromatin conformation I have been exploring a reconstitution approach which would be useful for establishing the molecular determinants of centromeric chromatin. The scheme is to take purified chromatin and to remove the weakly associated linker histone and non-histone proteins by passing

case then the occurrence of a monomer, dimer or trimer fragment is likely to be present at a similar frequency to a full-length fragment minus this detached portion which although difficult to quantitate, appears to be approximately correct (Figure 4.8F). It is difficult to know why the monomer, dimer or trimer fragments are prevalent, as a nick has the same probability of occurring throughout the full length of a long fragment.

As the sucrose gradients used for these studies were isokinetic, all the fragments (of the same conformation) will sediment at a rate directly proportional to their mass. For a linear-size distribution of fragments this would manifest itself as a straight line in the graphs of DNA size versus relative sedimentation (fraction number). However, the profile for the sedimentation of bulk chromatin in almost all the experiments (see also Chapter 3) tends to curve upwards (most clearly seen in the restriction digests e.g. Figure 4.8G), suggesting that longer chromatin fragments are more flexible than shorter fragments. This is in contrast to the profiles for the sedimentation of the satellite chromatin which are almost always linear. The simplest interpretation of this result is that as for a long stiff rod the satellite fragment can be longer before it will start to bend whilst, in analogy to a more flexible rod, the bulk chromatin fragments will start to bend if they have a shorter length. Bulk chromatin may therefore possess discontinuities and a greater inherent flexibility (Figure 4.1A) shown by a non-linear profile for its sedimentation, whereas satellite chromatin is more rigid and stable (Figure 4.1C) as shown by a linear profile in these gradients.

My preliminary results for the sedimentation of human centromeric chromatin are similar to results obtained for the mouse satellite. However, as mentioned earlier it is difficult for me to determine what sequences my probe is identifying as I do not have a panel of single chromosome genomic DNAs, and it possible that the probe is only picking up a subset of chromosomes. Only one human centromere (chromosome 10) has been characterised in detail (Jackson et al., 1999) and from this it is apparent that this human centromere is composed of a complex array of satellite sequences including α-satellite, classical satellite 2 and satellite 3 (Figure 1.10). To further explore the chromatin conformation of centromeres it will be necessary to analyse

the chromatin over a DNA cellulose column (Allan et al., 1981; Thoma and Koller, 1981). This will cause the 30nm chromatin fibre to unfold and will allow both bulk and satellite chromatin to sediment in a sucrose gradient at the same rate. By adding back purified proteins to the stripped chromatin it will be possible to determine the components required to create this altered chromatin structure. To date, I have stripped chromatin by passing it over a DNA cellulose column but I have found that satellite chromatin tends to aggregate whereas the bulk chromatin is far less susceptible to aggregation. This result is a further indication of the altered conformation of satellite chromatin and is possibly to be expected if the chromatin forms more regular and compact fibres. The next step in these experiments will be to optimise the conditions for this reconstitution approach.

My results suggest that centromeric heterochromatin is more rigid and compact than bulk chromatin which may be critical for the formation of a stable centromere. One biological reason for this might be that for the separation of sister chromosomes, which is quite a violent event as evidenced by dicentric chromosomes tearing themselves apart, a rigid underlying chromatin structure is necessary to form a sufficiently strong base for the kinetochore plate to attach to (Figure 4.11A). Alternatively, during meiosis I and mitosis, when the sister chromatids pair, having two rigid regions coming towards each other will increase the surface area between the two chromosomes and facilitate their interaction (Figure 4.11B). However, as the number of stable neocentromeres which have been identified increases it becomes harder to explain a role for these underlying centromeric sequences (Eichler, 1999). One argument which has been put forward is that the role of the satellite sequences may be to facilitate, but not be necessary for, the late replication of the centromeric region. This would determine the location of CENP-A and the subsequent position of the centromere (Csink and Henikoff, 1998). However, in some fashion neocentromeres are able to circumvent this requirement with sufficient efficiency to pass unaided through meiosis and mitosis (Tyler-Smith et al., 1999).

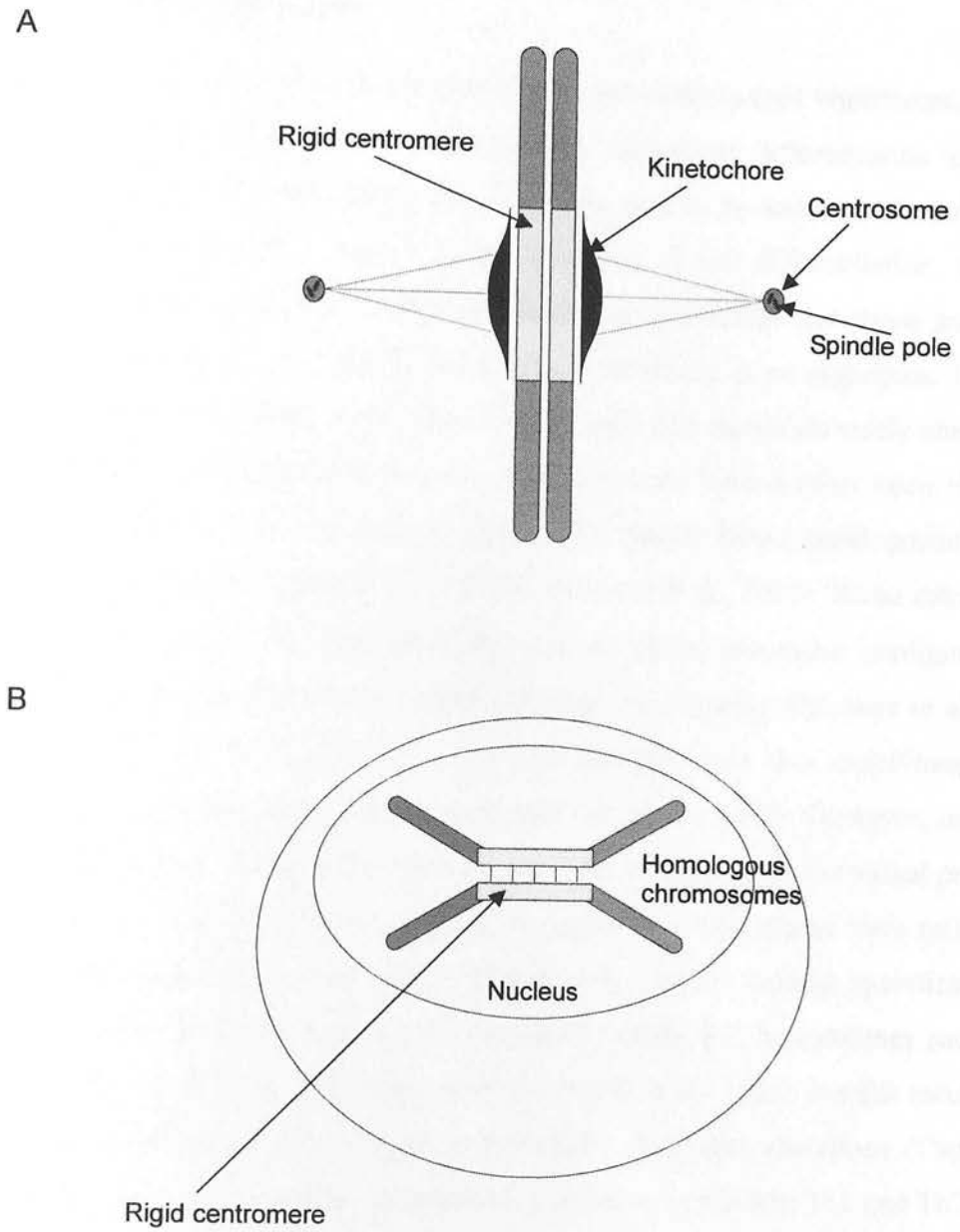


Figure 4.11 Models for the role of a rigid centromere. (A) During meiosis the sister chromatids move towards the spindle poles. The presence of a rigid centromere might be important for generating a stable contact to the kinetochore. (B) In mitosis and meiosis I the chromosomes pair for either duplication or homologous recombination. A rigid centromere might be required for ensuring the maximum surface area is available for alignment of the centromeres. After the centromeres have come together this may act as a nucleation point for the joining of the chromosome arms.

5.1 Introduction

The mechanisms responsible for maintaining stem cells in their undifferentiated state and the processes which occur during their subsequent differentiation are poorly characterised. One area which has not been studied in detail is the role of the underlying chromatin structure in the regulation of cell differentiation. Localised changes in the chromatin architecture have been identified and these presumably facilitate the access of specific transcription factors to gene sequences. However, alterations at the higher order fibre, domain level and above are rarely studied. For many years the supposition has been that stem cells have a more open chromatin architecture which is progressively repressed or altered during development (Caplan and Ordahl, 1978; Agarwal et al., 1999; Belmont et al., 1999). Some data suggest this: for example, the β-globin LCR is in an active chromatin configuration in hematopoietic progenitor cells prior to lineage specification (Jimenez et al., 1992) and studies on the erythroid and myeloid lineages show that multilineage locus activation precedes unilineage commitment (Hu et al., 1997). However, other data shows that multilineage gene expression can not be detected in individual progenitor cells suggesting that lineage specification occurs in a subclass of stem cells before differentiation (Brady et al., 1995; Berardi et al., 1995). Lineage specification has been shown to be triggered by the priming of specific loci by cytokines such as G-CSF (Heyworth et al., 1990) and EPO (Heyworth et al., 1995), but the mechanisms of commitment probably also involve multiple chromatin alterations (Caplan and Ordahl, 1978). For example, differentiation of naïve T cells into Th1 and Th2 cells is associated with long-range changes in the chromatin structure of effector cytokine genes (Agarwal and Rao, 1998). Also, studies in T-cells have shown that during activation there is a relocation of specific genes to heterochromatin suggesting that in this case, gene repression is mediated by their localisation to inactive chromatin domains (Brown et al., 1997a; Brown et al., 1999).

The chromatin fibre is frequently modified by histone acetylation, DNA methylation, or by altering linker histone subtypes. It has been shown that changing the level of

DNA methylation and histone acetylation alters the differentiation of cells (Taylor and Jones, 1979; Burns et al., 1988). During development, a number of different linker histones are expressed (Wiekowski et al., 1997; Clarke et al., 1998). These may alter the chromatin structure in a fashion that may aid or indeed facilitate cell differentiation (Steinbach et al., 1997).

Although my previous data have shown that the conformation of the higher order chromatin fibre does not change during the differentiation of stem cells, this work did suggest that there was an alteration in the nuclease accessibility of stem cell chromatin as compared to that of differentiated cells (Chapter 3). As it has been suggested that chromatin structure modulation may be important for stem cell differentiation, I thought it would be informative to alter the higher order chromatin fibre of mouse embryonic stem cells to ascertain whether this would change the differentiation potential of these cells.

Embryonic stem cells are derived directly from the inner cell mass (epiblast) of the pre-implantation mouse embryo (Evans and Kaufman, 1981; Martin, 1981; Brook and Gardner, 1997). These cells have the potential to differentiate into all other cell types including the germ line which permits germline transmission of the ES cell genotype and thereby provides a system for introducing predetermined modifications into experimental animals (Robertson, 1986; Robertson et al., 1986). ES cells are unique in retaining their capacity for multilineage differentiation both *in vitro* (Doetschman et al., 1985; Weiss and Orkin, 1996) and *in vivo* (Bradley et al., 1984; Beddington and Robertson, 1989; Nagy et al., 1990; Nagy et al., 1993). Therefore, they are ideally suited to being used as a model to identify and define the regulatory factors responsible for maintaining stem cells in their undifferentiated state and for the establishment of differentiated lineages. ES cells have the capacity both to produce identical daughter cells (self-renewal) and to produce daughters that are fated to differentiate. Thus ES cells are the only non-transformed mammalian stem cells that can be continuously propagated *in vitro*. ES cell self-renewal is sustained by the cytokine LIF (Smith et al., 1988; Williams et al., 1988). The actions of LIF are mediated via heterodimerisation of the low affinity LIF receptor and GP130 (Taga,

1997) which activates the STAT3 pathway (Niwa et al., 1998; Raz et al., 1999; Matsuda et al., 1999). ES cells can be differentiated into various lineages (Keller, 1995; Weiss and Orkin, 1996), including hematopoietic (Weiss, 1997), cardiac (Wobus and Guan, 1998), lymphocyte (Cho et al., 1999), adipocyte (Dani et al., 1997) and neuronal cell types (Fraichard et al., 1995; Bain and Gottlieb, 1998; Brustle et al., 1999). Some of these have been transferred directly into adult mice with evidence of functional, albeit transient, engraftment (Hole et al., 1996; Klug et al., 1996; Deacon et al., 1998) and may offer a source of cells for future transplantation (Dinsmore et al., 1996; Kind and Colman, 1999; Svendsen and Smith, 1999).

Previously, researchers have modulated the differentiation potential of cells by the addition of the histone deacetylase inhibitors trichostatin A or sodium butyrate, and DNA methyltransferase inhibitors such as 5-azacytidine (Taylor and Jones, 1979; Burns et al., 1988; Glauber et al., 1991). It has even been suggested that treatment of cells with 5-azacytidine causes a reversion to a more pluripotential state, possibly by altering the expression of genes through changes in the chromatin fibre (Taylor and Jones, 1979). One criticism of these approaches is that the target specificity of these agents is very wide and therefore the effects seen may be due to non-specific, non-chromatin mediated events. In addition these inhibitors may be targetting specific regulatory genes and either activating them by hyper-acetylation or inhibiting them by hyper-methylation rather than having a substantial effect on the global conformation of the chromatin fibre.

Histone H5 is normally expressed in nucleated chicken erythrocytes where it represses gene transcription and essentially deactivates all but a few housekeeping genes. It is expressed together with the other histones in dividing erythrocyte cell precursors, and is synthesised at a high level during erythrocyte maturation in a stage-specific manner (Affolter et al., 1987), most particularly as the cells enter the final stages of differentiation (Rousseau et al., 1993). This results in the accumulation of H5 in the chromatin of the mature cell, resulting in the partial replacement of histone H1 (Ruiz-Carrillo et al., 1974). H5 expression is correlated

with chromatin condensation and a decrease in replication and transcription (Appels and Wells, 1972). Induced expression of a transfected H5 gene in rat sarcoma cells to levels similar to those in mature chicken erythrocytes causes reversible inhibition of DNA replication (Sun et al., 1989). This replacement of H1 by H5 was found to enhance the stability of the chromatin fibre without altering its conformation (Sun et al., 1990c). Microinjection of H5 into rat L6 myoblasts shows that H5 is concentrated in the nucleus, reducing transcription and replication substantially, with the chromatin of the injected cells becoming densely packaged (Bergman et al., 1988). Ectopic expression of H5 only mediates an effect when the protein is dephosphorylated (Aubert et al., 1991) as in avian erythrocytes where it is dephosphorylated when the cell matures (Ruiz-Carrillo et al., 1976; Sung, 1977). A structurally and perhaps functionally related linker histone, H1°, has been immunolocalised to condensed regions of chromatin when MEL cells are committed to differentiate (Osborne and Chabanas, 1984), although H1° is by no means excluded from active chromatin (Gorka et al., 1993). The level of H1° accumulates postmitotically (Gjerset et al., 1982) and increases in cells that are either terminally differentiated or have a lower rate of cell division (Pehrson and Cole, 1980; Chabanas et al., 1983). Some H1° expression is modulated by acetylation, with trichostatin A efficiently inducing H1° gene expression (Girardot et al., 1994). However, it does not appear that H1° is required for normal cellular function as H1° knock-out mice seem to be healthy (Sirotkin et al., 1995) and transgenic mice over-expressing human H1° exhibit a normal phenotype (Tonjes et al., 1997). The function of H1°, and thereby its functional relationship with H5, remains unknown.

Green fluorescent protein (GFP) from the jellyfish *Aequorea victoria* is a fluorescent molecule suitable for monitoring gene expression and localisation in eukaryotic cells (Chalfie et al., 1994; Cubitt et al., 1995). The wild-type protein has been modified to give enhanced spectral properties (Cormack et al., 1996) including FACS-optimisation and more intense fluorescence, and recently a variety of differently coloured GFP derivatives have been made. A histone H2B-GFP fusion has been stably expressed in HeLa cells showing that the histone octamer can accommodate

the H2B-GFP fusion protein (Kanda et al., 1998). In this case, the construct was used as an *in vivo* tag for fluorescently staining chromatin to monitor minute chromosomes. In addition, in yeast, an H1-GFP fusion has been made to investigate the localisation of the putative yeast H1 protein to the nucleus (Ushinsky et al., 1997). Also, it has been shown that GFP can be fused to Polycomb, a chromatin associated protein responsible for repressing homeotic genes. The fusion protein was used to monitor the localisation of Polycomb during the development of *Drosophila* embryos, and although it was able to bind chromosomal DNA, it was non-functional (Dietzel et al., 1999).

With a view to elucidating the contribution of chromatin structure to the regulation of stem cell differentiation, I have expressed histone H5 in ES cells. Histone H5 would be expected to condense the chromatin fibre and perhaps thereby alter the phenotype of the cells, it being generally thought that the chromatin fibre is naturally condensed during lineage selection.

5.2 NIH3T3 cells can express histone H5

Linker histone variants are expressed in a developmental stage specific manner and have been associated with alterations in cell differentiation potential. We were interested in expressing histone H5 in ES cells to determine whether this may alter the cells differentiation or stem cell renewal capacity, and to further evaluate any special features of H5. As a preliminary step, I transfected histone H5 into mouse NIH3T3 embryonic fibroblast cells to confirm that the expression constructs functioned correctly and to determine whether the expression of H5 induced any overt phenotype. The NIH3T3 cell type has been maintained in culture for many years; they are easy to grow, are readily transfected, and are more resilient to the ectopic expression of proteins than some other cell lines, making them ideally suited to this type of expression study.

To make the expression constructs, the chicken histone H5 gene was subcloned from plasmid pBRH52.4 (Plasmid #16; a gift from J. Dodgson) using NotI and XmnI. The NotI 5'-overhang was filled using Klenow DNA polymerase and the fragment was

subcloned into two expression vectors at a blunted XhoI site. The two constructs (described in detail in section 5.3) are identical except that one (pPHCAGGS-BstXI; Figure 5.5C) has a ubiquitously strong chicken β-actin (CAG) promoter (Niwa et al., 1991), and the other (pPHTK-BstXI; Figure 5.5D) has a weak thymidine kinase (TK) promoter (Thomas and Capecchi, 1987). The resulting plasmids are called pPHCAGGS-H5NX and pPHTK-H5NX, respectively. NIH3T3 cells were transiently transfected with these constructs using lipofection (Section 2.6.7). Nuclei were prepared from the cells after 24 hours (Section 2.7.3) and the protein content analysed on a 15% SDS-PAGE gel (Section 2.10.2; Figure 5.1A). The gel was western blotted (Section 2.10.3) with a monoclonal antibody, 4C6, against histone H5 (Mendelson and Bustin, 1984) (Figure 5.1B).

The transfected cells expressed histone H5 at the correct molecular weight. The stronger CAG promoter construct gave a much higher level of expression than the TK promoter construct. The blot also shows that the monoclonal antibody does not cross-react with any endogenous protein found in the NIH3T3 cells, in contrast to polyclonal antibodies which frequently cross-react with H1° (Allan et al., 1982b).

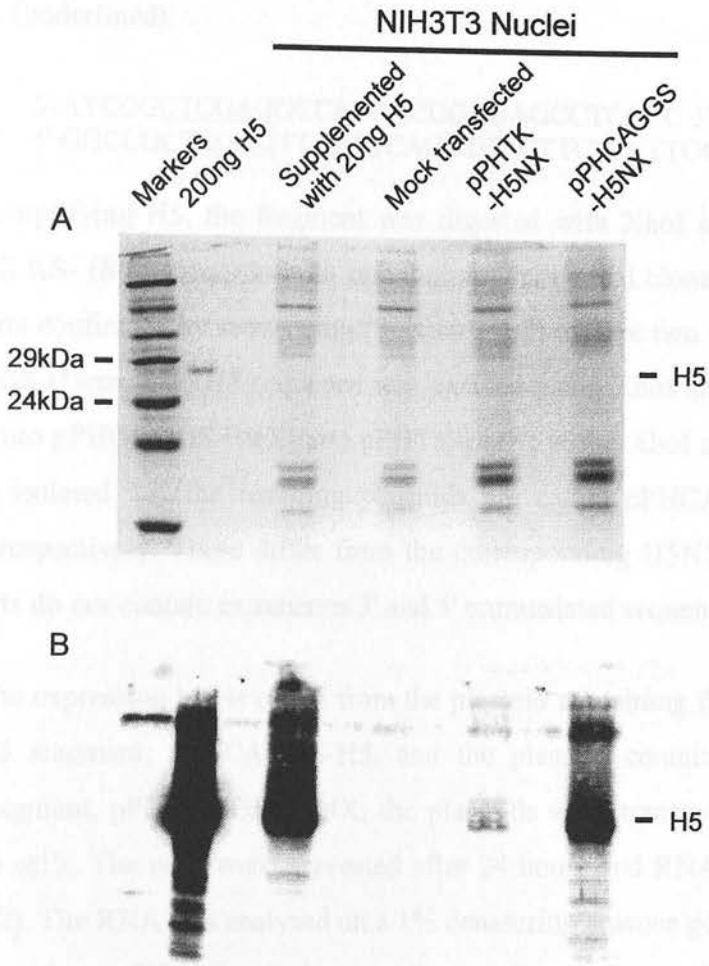


Figure 5.1 NIH3T3 cells express histone H5. NIH3T3 cells were transfected with plasmids which express H5. After 24 hours the cells were harvested and nuclei prepared. (A) Nuclear proteins were analysed on a 15% SDS-PAGE gel and western blotted with an antibody against H5 (B). 5μg of nuclei were loaded in each lane. 200ng H5 is equivalent to 1μg nuclei.

Although the original construct expressed H5 correctly there was the possibility that the NotI/XmnI H5 fragment, which also contains extra 3'- and 5'- untranslated sequences, may encode elements which could interfere with the stable expression of H5 in further experiments. Therefore, the H5 coding sequence was PCR amplified (Section 2.4.7) using the following gene-specific primers that carry XhoI restriction enzyme sites (underlined).

H5A.for: 5'-ATCCGCTCGAGGCCATGACGGAGAGCCTGGTC-3'
H5A.bak: 5'-GGCCGCTCGAGTTACTTCAGCTCACTTCTTCTTGGGCGATTT-3'

After PCR amplifying H5, the fragment was digested with XhoI and ligated into pBluescript II KS- (Stratagene) in both orientations. Individual clones were isolated and the inserts confirmed by sequencing (Section 2.4.8) to give two plasmids: pBS-H5for and pBS-H5rev. The H5 sequence was excised using XhoI and the fragment was cloned into pPHCAGGS-BstXI and pPHTK-BstXI at the XhoI sites. Individual clones were isolated and the resulting plasmids are called pPHCAGGS-H5, and pPHTK-H5 respectively. These differ from the corresponding H5NX constructs in that the inserts do not contain extraneous 3' and 5' untranslated sequences.

To analyse the expression levels of H5 from the plasmid containing the shorter PCR amplified H5 fragment, pPHCAGGS-H5, and the plasmid containing the larger restriction fragment, pPHCAGGS-H5NX, the plasmids were transiently transfected into NIH3T3 cells. The cells were harvested after 24 hours and RNA was prepared (Section 2.7.2). The RNA was analysed on a 1% denaturing agarose gel and Northern blotted using probes to H5 and a single copy ribosomal control gene, S26 (Vincent et al., 1993) (Figure 5.2A). It appears that the shorter PCR-generated H5 fragment is expressed slightly more efficiently than the longer H5 restriction fragment. However, as the control probe (S26) is against a cellular protein it does not control for the plasmid copy number. From the blot it is clear that the H5 restriction fragment construct (pPHCAGGS-H5NX) gives a longer transcript due to its extended 5' and 3' untranslated regions. For all future experiments the shorter, and more highly expressed PCR amplified H5 construct was used.

To estimate the level of H5 expression in the NIH3T3 cells, nuclei were prepared from transiently transfected cells. Known amounts of nuclei were fractionated on a SDS-PAGE gel and were western blotted using the monoclonal antibody against H5 (Figure 5.2B). The expression level of H5 from the pPHCAGGS-H5 construct in NIH3T3 cells is approximately 30% of that found in chicken erythrocytes, taking into account the transfection efficiency, and in agreement with the results shown in Figure 5.1.

To determine whether H5 could be stably expressed in NIH3T3 cells without altering the cell's phenotype I transfected cells with the two H5 expressing constructs pPHCAGGS-H5 and pPHTK-H5. After 24 hours the cells were replated at a clonal density in duplicate and those which had stably incorporated the plasmid using hygromycin were selected (Table 2.2). After 8 days selection with hygromycin all the untransfected cells had died, and the plates were stained with Leishman's fix and stain to visualise the colonies. From the results (Figure 5.3A), it appears that H5 does not notably affect the survival of the transfected cells. The remaining transfected cells were passaged and expanded. Nuclei were prepared from the cells and analysed for H5 expression by western blotting (Figure 5.3C). Although the cells did express histone H5, it was at a level much lower than that found in chicken erythrocytes suggesting that either a higher level of H5 expression was detrimental to the cells as has been shown previously (Bergman et al., 1988; Aubert et al., 1991) or that the promoter driving the construct was in some down-regulated in the stable transfectants. In duplicate experiments it was never possible to see expression of H5 from the TK promoter construct, presumably because it suffered from position dependent silencing to a greater extent than the CAG promoter construct, and because the sensitivity of the 4C6 antibody was insufficient to detect the small amount which was being expressed.

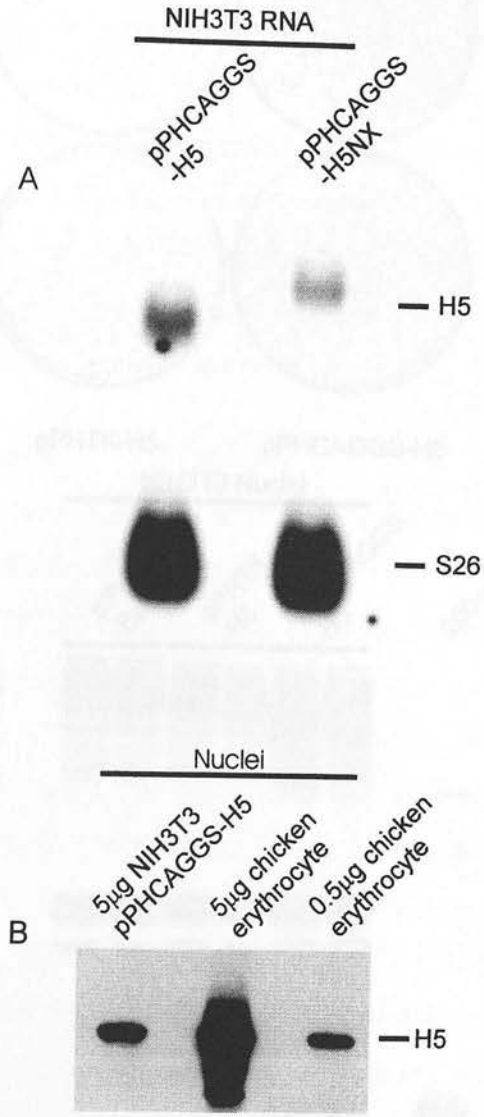


Figure 5.2 NIH3T3 cells express a short PCR generated H5 fragment slightly more efficiently than a longer H5 restriction fragment. NIH3T3 cells were transiently transfected with plasmids expressing H5. (A) After 24 hours RNA was prepared from the cells and fractionated on an agarose gel. The gel was northern blotted and hybridised to an H5 or S26 control probe. (B) After 24 hours nuclei were prepared from the cells and analysed by SDS-PAGE and western blotting with an H5 antibody. In this experiment the level of H5 expression was found to be approximately 30% the level of H5 normally found in chicken erythrocytes, having accounted for the transfection efficiency.

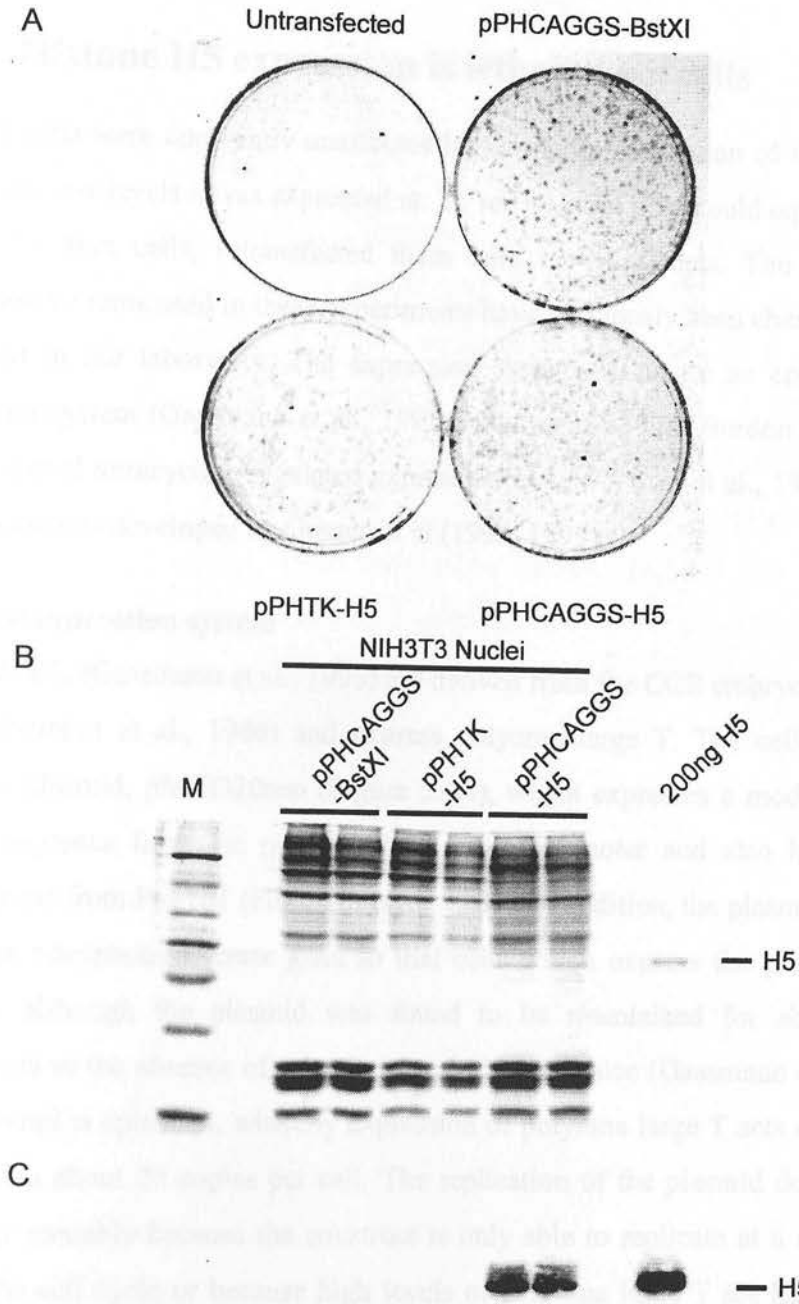


Figure 5.3 Stable H5 expression in NIH3T3 cells does not appear to inhibit cell survival, but the cells do express H5 at a much lower level than that found in chicken erythrocytes. (A) NIH3T3 cells were transfected with a control vector (no insert) and H5 expressing plasmids. After 24 hours the cells were replated at a clonal density in duplicate and selected in the presence of hygromycin. After 8 days selection one set of plates was stained with Leishman's fix and stain to visualise surviving colonies. The remaining plates were passaged and expanded. (B, C) Nuclei were prepared from the cells and analysed by SDS-PAGE and western blotting with an H5 antibody. 5μg nuclei were loaded per lane in duplicate. 200ng H5 is equivalent to 1μg chicken erythrocyte nuclei.

5.3 Histone H5 expression is lethal to ES cells

NIH3T3 cells were apparently unaffected by the stable expression of histone H5, at least at the low levels it was expressed at. To see whether this would equally apply to embryonic stem cells, I transfected them with my constructs. The two ES cell expression systems used in these experiments have previously been characterised and optimised in our laboratory. The expression systems used are an episomal based expression system (Gassmann et al., 1995; Niwa et al., 1998; Burdon et al., 1999), and a targeted tetracycline-regulated expression system (Niwa et al., 1998) based on the components developed by Gossen et al (1992; 1995).

Episomal expression system

MG1.19 cells (Gassmann et al., 1995) are derived from the CCE embryonic stem cell line (Robertson et al., 1986) and express polyoma large T. The cells contain an episomal plasmid, pMGD20neo (Figure 5.4A), which expresses a modified large T coding sequence from the polyoma late region promoter and also has a mutant enhancer-ori from PyF101 (Fujimura et al., 1981). In addition, the plasmid contains a neomycin phosphotransferase gene so that cells which express the plasmid can be selected, although the plasmid was found to be maintained for about 50 cell generations in the absence of selection for G418 resistance (Gassmann et al., 1995). This plasmid is episomal, whereby expression of polyoma large T acts on the origin to maintain about 20 copies per cell. The replication of the plasmid does not 'run-away', presumably because the construct is only able to replicate at a certain point during the cell cycle or because high levels of polyoma large T are harmful to the cells.

This cell line is able to support a second plasmid containing a polyoma ori in a similar episomal state. The expression plasmids used in conjunction with this cell line carry an expression cassette, a polyoma origin, a selectable marker, and a cloning site. For this study the expression cassette used was a strong cytomegalovirus enhancer, a chicken β-actin promoter and first non-coding exon (Niwa et al., 1991) with a chicken β-actin intron, and a rabbit β-globin splice acceptor and poly-A signal

(CAG cassette; Figure 5.4B), or a TK promoter with a rabbit β-globin poly-A tail. Similar constructs have been made in the laboratory which have other promoters, including PGK (phosphoglycerate kinase), hBA (human β-actin), and SV40p/e (SV40 promoter and enhancer) (Hitoshi Niwa, unpublished). The selectable marker carried on these expression plasmids was the *hph* gene which encodes for hygromycin phosphotransferase which can be selected for using hygromycin B. The cloning site has XhoI and BstXI restriction sites. The two BstXI sites are derived from the pCDM8 (Invitrogen) polylinker and are designed so that once the vector is cut it is unable to re-ligate in the absence of an insert. The only disadvantage of this approach is that suitable adapters have to be ligated onto the ends of the insert (Section 2.4.6) in a process which is quite inefficient. The complete expression plasmids are referred to as pPHCAGGS-BstXI (Figure 5.4C), and pPHTK-BstXI (Figure 5.4D).

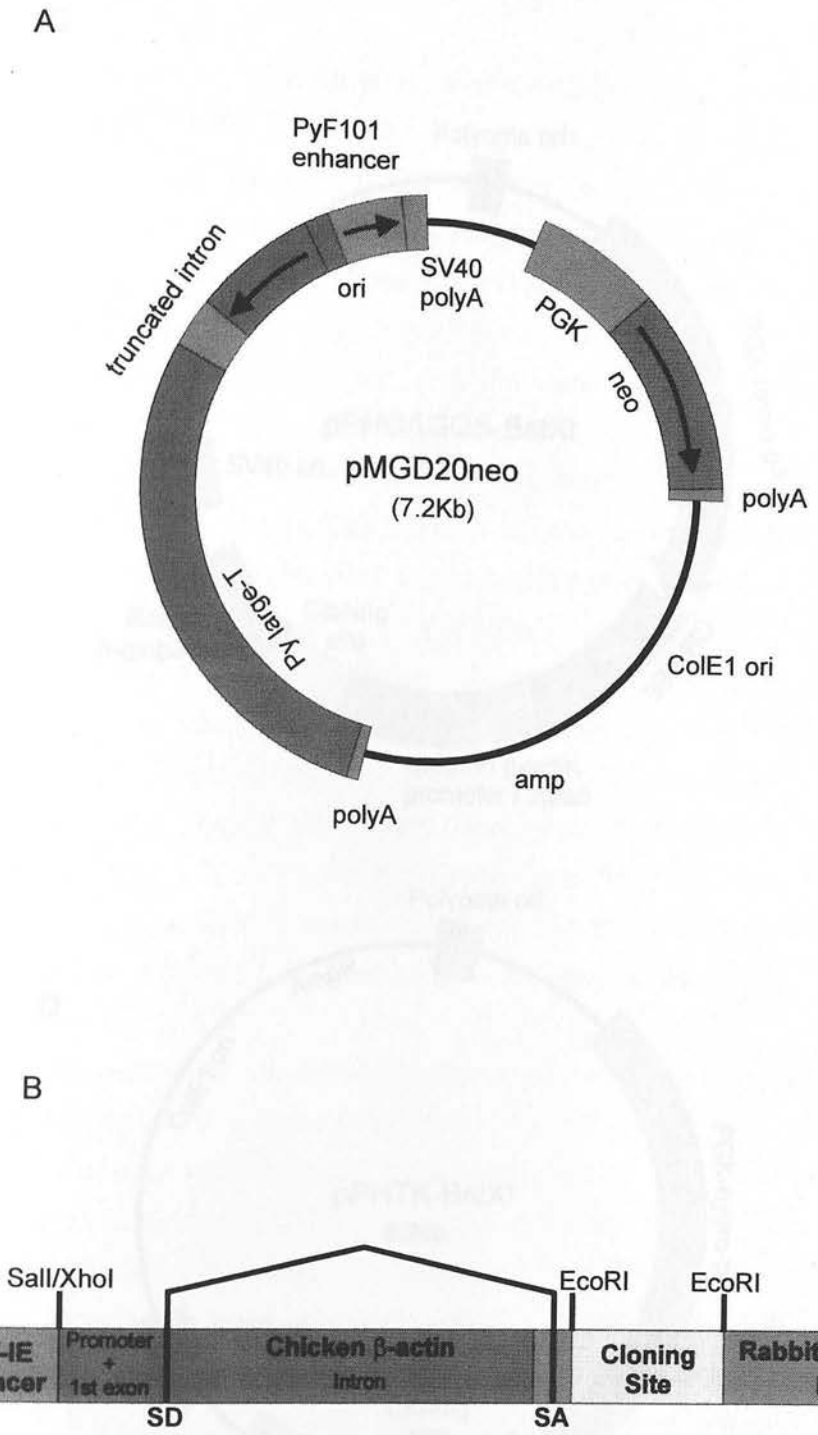
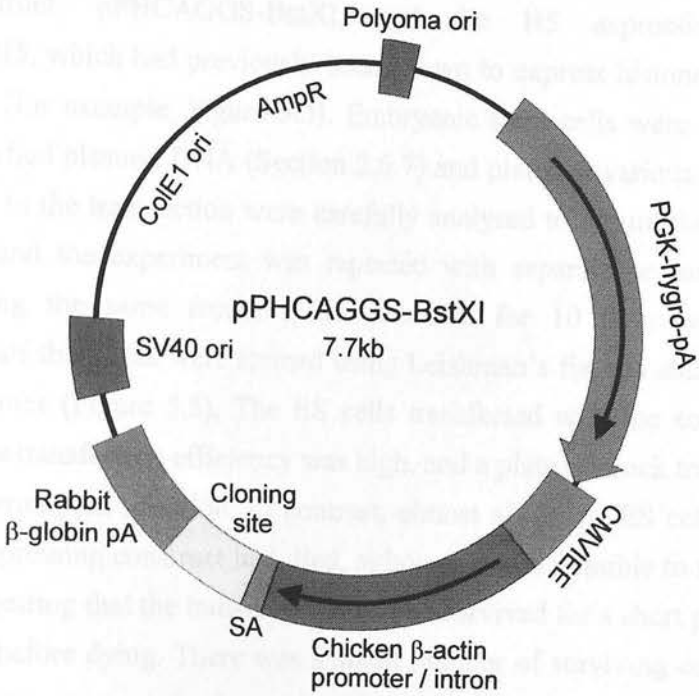


Figure 5.4 Components of the episomal plasmid expression system. (A) MG1.19 ES cells carry the plasmid pMGD20neo which expresses polyoma large T. This plasmid contains a polyoma origin allowing it to be maintained as an episome. (B) The CAG cassette provides high level ES cell expression. (cont.)

C



D

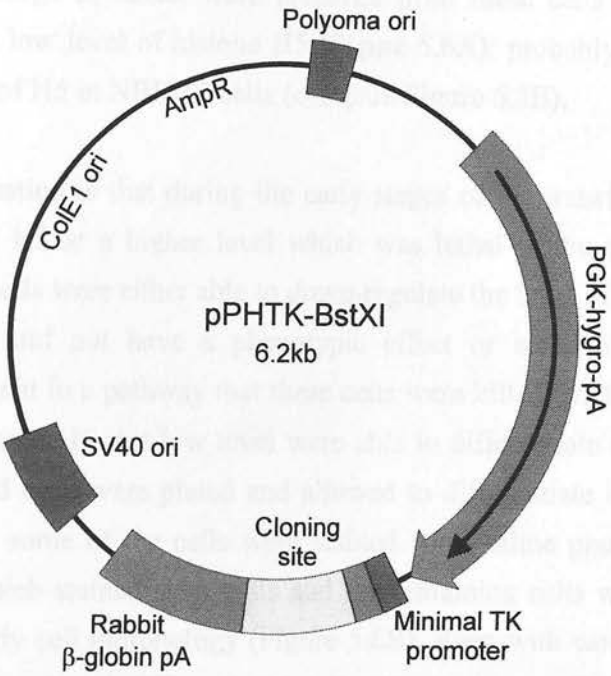


Figure 5.4 (cont.) (C, D) The two expression vectors, pPHCAGGS-BstXI and pPHTK-BstXI, can be maintained episomally in MG1.19 cells. They encode a hygromycin selection marker and a suitable cloning site for expressing cDNAs. The CAG vector expresses cDNAs at about 20-times the level of the TK vector (H. Niwa, unpublished).

In the initial experiment, MG1.19 embryonic stem cells were transfected with the control construct, pPHCAGGS-BstXI, and the H5 expressing construct, pPHCAGGS-H5, which had previously been shown to express histone H5 in mouse NIH3T3 cells (for example, Figure 5.3). Embryonic stem cells were electroporated with 20μg purified plasmid DNA (Section 2.6.7) and plated at various densities. The plasmids used in the transfection were carefully analysed to ensure they were of the same quality and the experiment was repeated with separate preparations of the plasmids giving the same result. After selection for 10 days with 100μg/ml hygromycin, half the plates were stained using Leishman's fix and stain to show the surviving colonies (Figure 5.5). The ES cells transfected with the control plasmid showed that the transfection efficiency was high, and a plate of mock transfectants all died under hygromycin selection. In contrast, almost all of the ES cells transfected with the H5 expressing construct had died, although it was possible to see clumps of dead cells suggesting that the transfected cells had survived for a short period of time and expanded before dying. There was a small number of surviving colonies which appeared to have a normal phenotype; these were maintained in culture and expanded. At passage 3, nuclei were prepared from these cells and found to be expressing a very low level of histone H5 (Figure 5.6A); probably less than for the stable expression of H5 in NIH3T3 cells (compare Figure 5.3B).

A possible explanation is that during the early stages of the transfection most cells expressed histone H5 at a higher level which was lethal to the cells. However, a small fraction of cells were either able to down-regulate the level of H5 expression to a point where it did not have a phenotypic effect or some of the cells were functionally deficient in a pathway that these cells were killed by. To ensure that the cells which expressed H5 at a low level were able to differentiate correctly, control and H5 transfected cells were plated and allowed to differentiate in the absence of LIF. After 4 days some of the cells were stained for alkaline phosphatase activity (Section 2.6.4) which stained stem cells and the remaining cells were stained with Leishman's to study cell morphology (Figure 5.6B). Even with careful examination of the undifferentiated and differentiated cells it was not possible to identify any differences between the control and H5 expressing cells. This suggests that this level of H5 expression is unable to exert a phenotypic effect or this sub-population of cells has adapted to a phenotypic alteration which may have occurred due to the expression of H5.

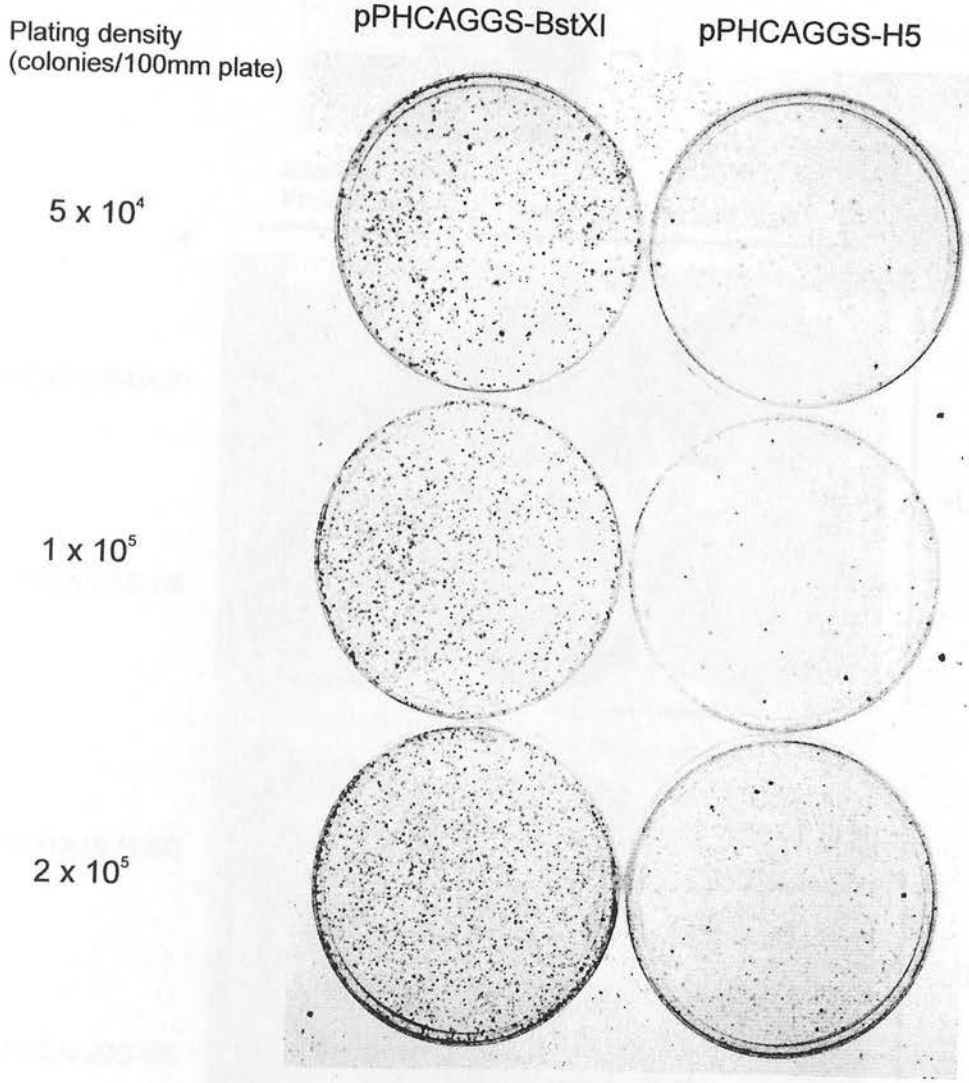


Figure 5.5 Histone H5 expression is highly toxic to ES cells. MG1.19 cells were transfected with an H5 expressing plasmid and a control vector (no insert). The cells were plated at various densities in triplicate and after 24 hours were selected with hygromycin. After 10 days, two sets of plates were stained using Leishman's stain to allow visualisation of the colonies. Representative plates are shown.

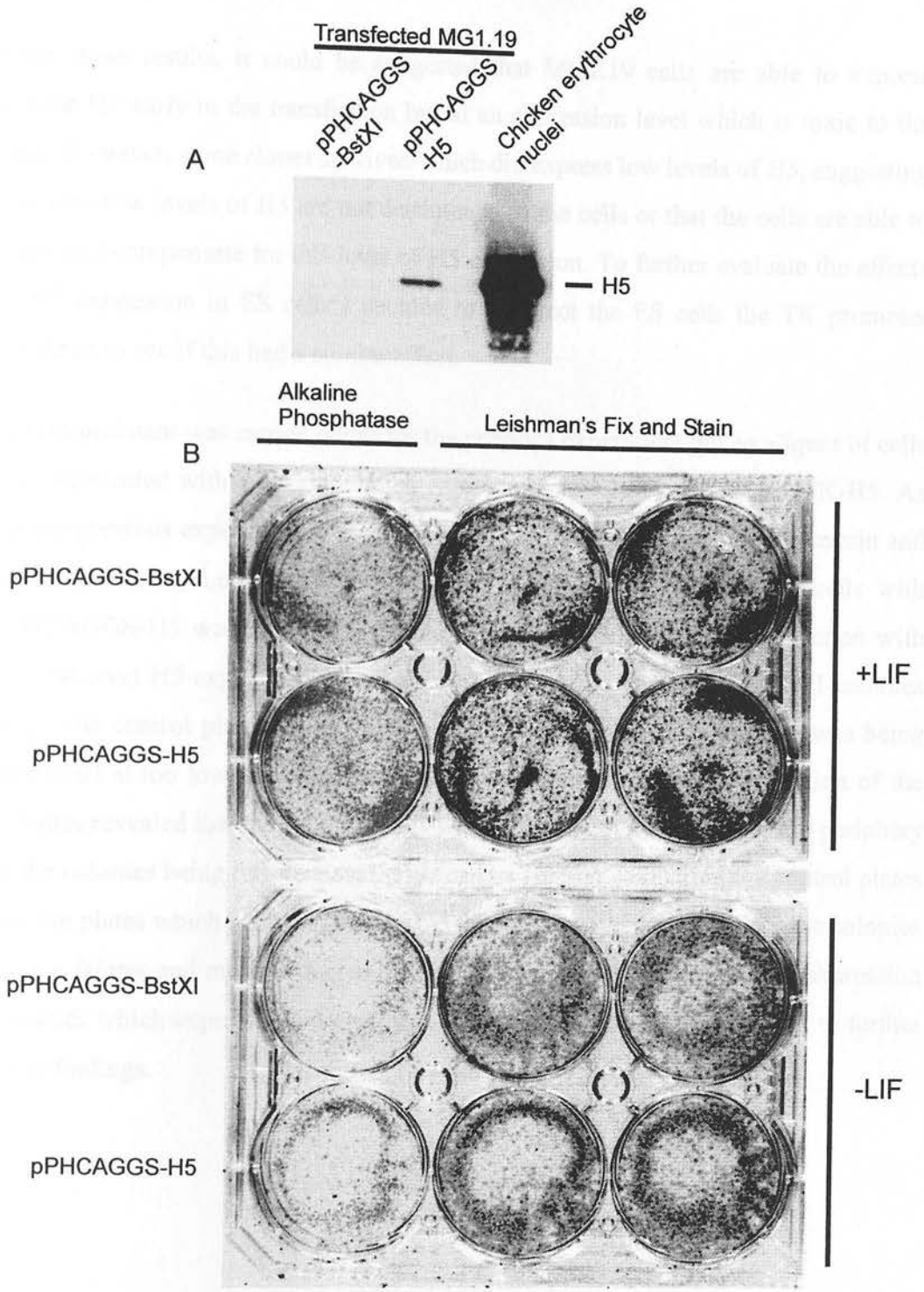


Figure 5.6 Histone H5 can be expressed in ES cells. The expression of H5 in ES cells caused most of the cells to die (Figure 5.5). The surviving cells were passaged and expanded. (A) H5 expression was analysed by preparing nuclei from the cells and western blotting with an antibody against H5. 5µg of nuclei were loaded in each lane. (B) To determine whether these transfected cells which expressed H5 differentiated correctly the cells were plated in the presence and absence of LIF. After 4 days the cells were either stained for alkaline phosphatase activity to identify ES cell colonies or with Leishman's fix and stain to show the cell morphology.

From these results, it could be suggested that MG1.19 cells are able to express histone H5 early in the transfection but at an expression level which is toxic to the cells. However, some clones survived which did express low levels of H5, suggesting that very low levels of H5 are not detrimental to the cells or that the cells are able to adapt and compensate for this level of H5 expression. To further evaluate the effects of H5 expression in ES cells I decided to transfect the ES cells the TK promoter construct to see if this had a similar effect.

This experiment was carried out as for the previous experiment but an aliquot of cells was transfected with a plasmid which expressed a low level of H5, pPHTK-H5. As for the previous experiment, the cells were selected for 8 days using hygromycin and then fixed using Leishman's fix and stain. Again, transfection of the cells with pPHCAGGS-H5 was found to be lethal (Figure 5.7). In contrast, transfection with the low level H5 expression vector resulted in a similar number of ES cell colonies as for the control plasmid. On initial examination, this suggests that H5 was being expressed at too low a level to have an effect. However, closer examination of the colonies revealed that their morphology was altered, with many cells at the periphery of the colonies being differentiated. This can be seen on comparing the control plates and the plates which had been transfected with a low level of H5, where the colonies have a fainter and more feathered outline (Figure 5.8). Time permitting, expression plasmids which express an intermediate level of H5 would have been used to further these findings.

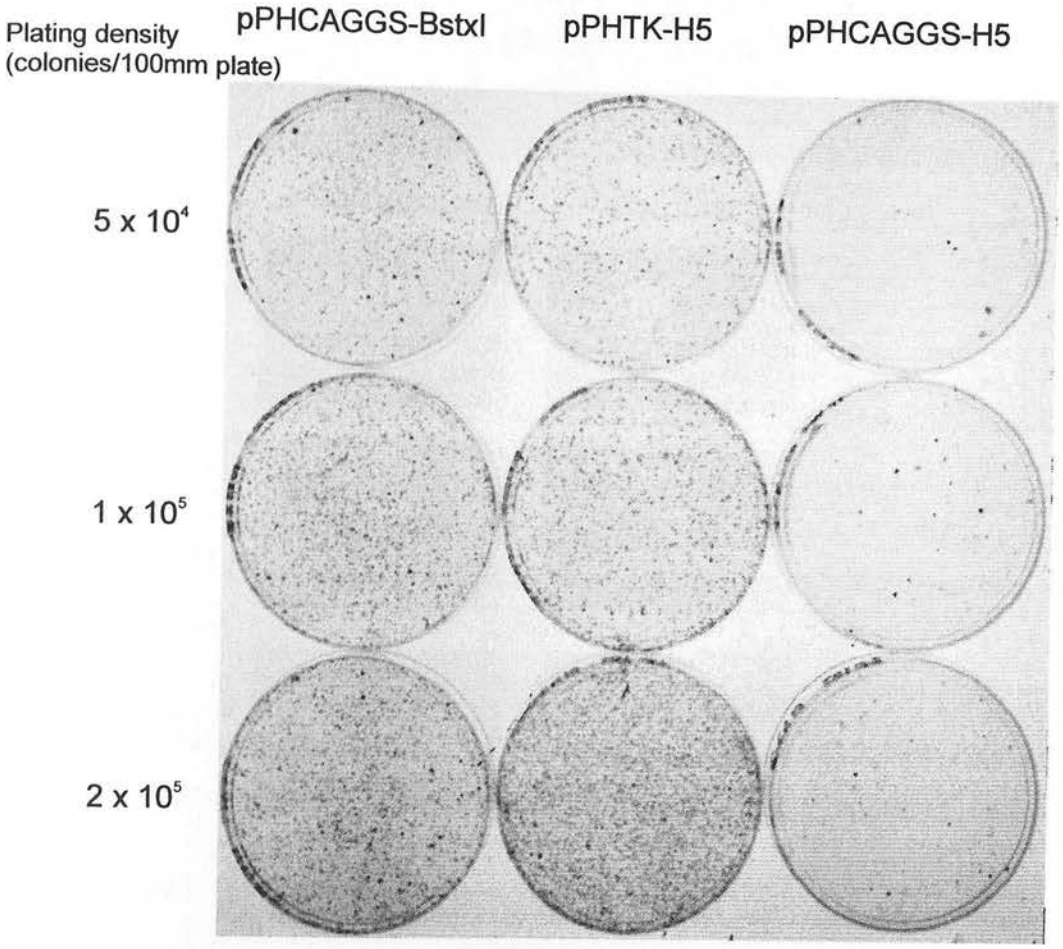


Figure 5.7 Low level H5 expression is not lethal to ES cells. MG1.19 ES cells were transfected with a control vector (no insert) or plasmids which expressed either a low level or a higher level of H5. After 8 days selection in hygromycin the cells were stained with Leishman's fix and stain to visualise the surviving colonies. As for the previous experiment (Figure 5.5) the cells transfected with a plasmid expressing a high level of H5 mostly died, whilst the cells transfected with a plasmid which expressed a lower level of H5 appeared to survive.

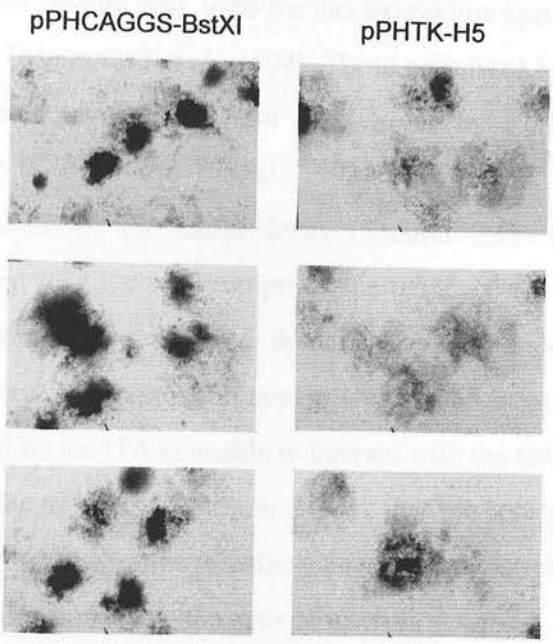


Figure 5.8 Low level H5 expression alters the ES cell phenotype. ES cells which had been transfected with a plasmid which expressed a low level of H5 were examined in detail. The ES cell colonies expressing H5 appeared to have a more feathered outline and contained more differentiated cells than the control cells.

Tetracycline-regulated expression system

A disadvantage with the colony forming assay system is that it is not possible to identify cells early in the transfection that are expressing histone H5, and therefore to assess whether this expression of H5 has a phenotypic effect. One solution is to use a tetracycline-regulated expression system which would allow a clonal population of cells to be expanded prior to the induction of H5 expression in all the cells synchronously.

The cell line, ZHTc6 (Figure 5.9), used for this expression system was derived from the CGR8 cell line (Mountford et al., 1994). These cells have been modified to carry a randomly integrated regulatable Oct-4 gene driven by a tetracycline-regulatable CMV*-1 promoter (Gossen and Bujard, 1992). In addition, the cells have been targeted to one allele of the Oct-4 locus (Section 3.1) with the tetracycline transactivator protein (tTA), which comprises the tetracycline repressor (tetR) fused to the herpes simplex virus C-terminal domain of VP16, under the control of the strong CAG promoter. The expression system is designed to work as follows: in the presence of tetracycline the tTA is unable to interact with the tetO operator sequences which are fused to the minimal CMV promoter so that the construct does not express the gene of interest. In the absence of tetracycline, the transactivator is able to bind to the operator sequences expressing the gene of interest.

This cell line took a considerable amount of time to develop as the transactivator protein is toxic to ES cells, but needs to be expressed at a sufficiently high level to ensure that expression from the promoter is regulatable. Once a suitable cell line had been screened and found to regulate the expression of Oct-4, it was thought easier to target into this locus than to randomly integrate the regulatable expression construct again, and have to screen for suitable clones. The approach I therefore used was to target the random integration of the regulatable Oct-4 gene in the ZHTc6 cells using a suitable replacement vector derived from the original targetting construct (SuperKO; Plasmid #24; Niwa et al., 1998) but which now expresses histone H5 (Figure 5.9).

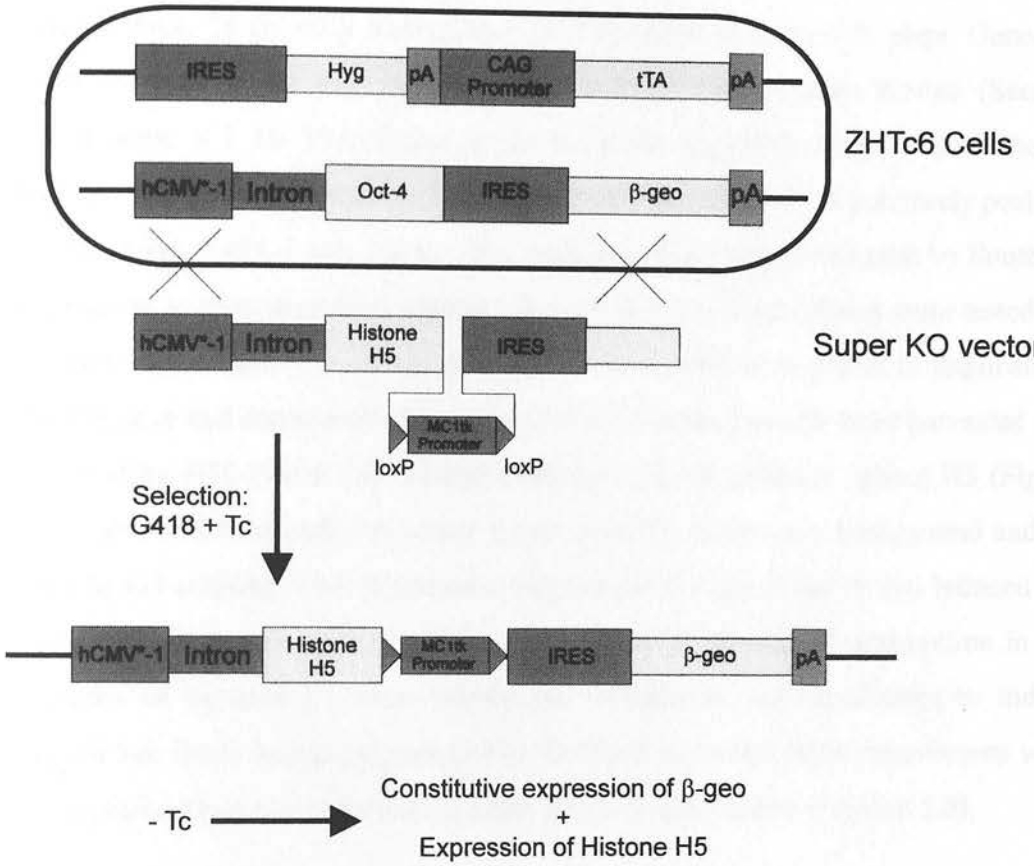


Figure 5.9 Targetting of ZHTc6 cells with an H5 replacement construct. ZHTc6 cells contain a tetracycline regulated promoter driving the Oct-4 cDNA. Targetting of these cells with the linearised Super KO replacement construct carrying H5 mediates the exchange of the Oct-4 cDNA with H5, by homologous recombination. Targetted clones can be selected in the presence of G418 and tetracycline. The targeted clones can be induced to express H5 by the withdrawal of tetracycline from the culture medium.

The H5 gene was subcloned from pBS-H5rev using XhoI and cloned into the SuperKO vector at the XhoI site and screened using colony lifts (Forster et al., 1990). The vector was linearised at the ScaI site for targetting. ES cells were transfected with 100μg linearised plasmid DNA and plated at a suitable density. The cells were selected in tetracycline (Table 2.2) to inhibit ZHTc6 β-geo expression, and G418 (Table 2.2) to select for targeted clones which will constitutively express β-geo. After 9 days, 48 colonies were picked and expanded in a 96-well plate. Genomic DNA (Section 2.7.1) was prepared from the cells and Southern blotted (Section 2.8.3) using a 2 kb HindIII/SacI probe from the SuperKO vector which should indicate positive colonies as a 3.2 kb fragment (Figure 5.10). Five putatively positive clones (1-B5, 1-C3, 1-D5, 2-B4, 2-B5) were expanded and re-screened by Southern blotting to confirm they were correct (data not shown). These clones were tested for H5 expression upon tetracycline withdrawal. The clones were plated in duplicate in the presence and absence of tetracycline. After 24 hours the cells were harvested and analysed by SDS-PAGE and western blotting using the antibody against H5 (Figure 5.11). Even with extended exposure it was possible to see only background and no specific H5 staining. This experiment suggested that none of the clones induced H5 expression. However, this may have been indicative of residual tetracycline in the medium, or because 24 hours tetracycline withdrawal was insufficient to induce expression. Even though targeted clones had been identified these experiments were not continued due to results from another series of experiments (Section 5.5).

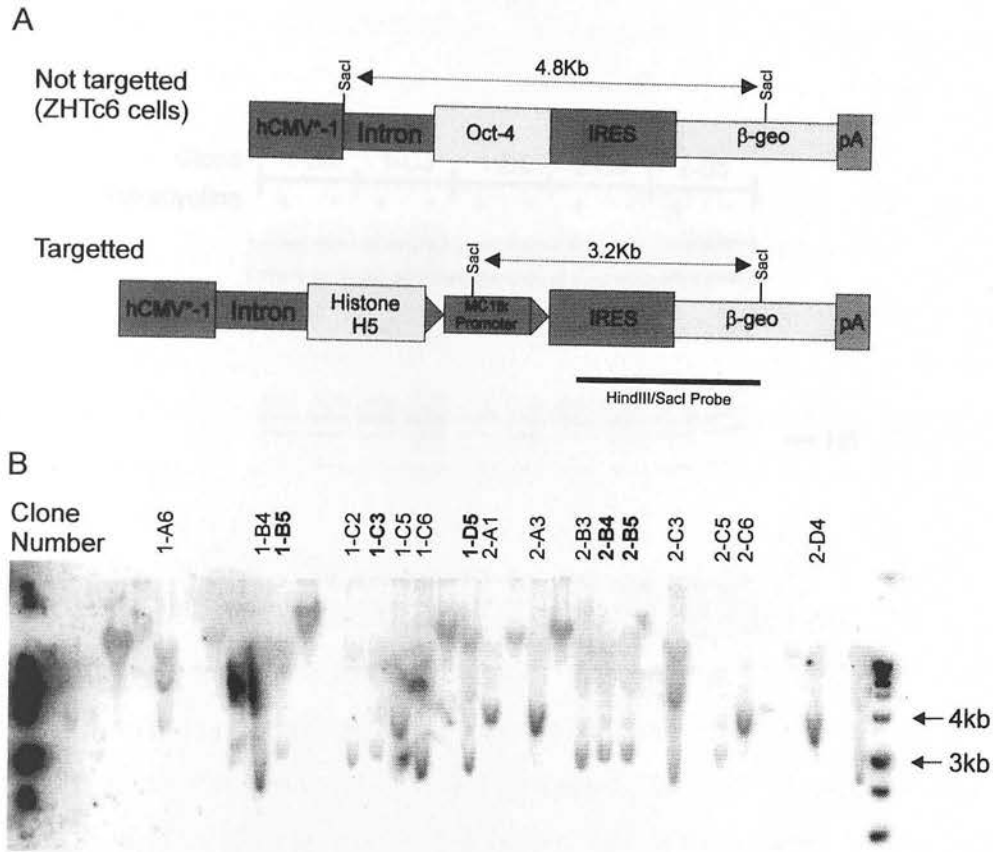


Figure 5.10 Targetting of the regulated H5 expression construct in ZHTc6 cells. ZHTc6 cells were transfected with linearised Super KO plasmid carrying H5. Cells were selected in G418 and tetracycline. After 9 days 48 clones were picked and genomic DNA was prepared. (A) The DNA was digested with SacI and fractionated on an agarose gel and Southern blotted to a HindIII/SacI probe. (B) Targeted clones will give a 3.2 kb band whilst untargeted ZHTc6 cells will give a 4.8 kb band.

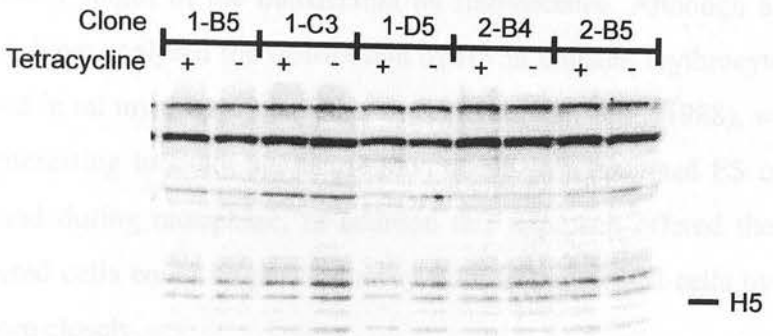


Figure 5.11 Tetracycline withdrawal did not regulate H5 expression. ZHTc6 cells targeted with H5 were grown in the presence and absence of tetracycline. After 24 hours total cell extracts were analysed by SDS-PAGE and western blotted with the H5 antibody. Even after a long exposure (as shown) it was not possible to see any specific H5 expression.

5.4 H5 fusion protein expression is localised away from centromeric heterochromatin

H5 expression in ES cells appears to have a phenotypic effect. As the colony forming assay is only suitable for studying clones after ten days selection and by then the most interesting phenotypic changes may have occurred, it was thought useful to transfect ES cells with a reporter construct so that transfected cells could be monitored immediately after transfection. The approach taken was to generate a GFP-H5 fusion protein (Gilbert and Allan, submitted) which could be monitored during the early stages of the transfection by fluorescence. Although a number of other studies have analysed the distribution of H5 in chicken erythrocytes (Mura et al., 1982) and in rat myoblasts injected with H5 (Bergman et al., 1988), we thought it would be interesting to study the distribution of H5 in transfected ES cells both in interphase and during metaphase. In addition this approach offered the possibility that transfected cells could be purified away from untransfected cells by FACS and analysed more closely.

Construction of GFP-H5 fusion protein

A large number of functional GFP fusion proteins have been generated (Cubitt et al., 1995) including two histone protein fusions (Ushinsky et al., 1997; Kanda et al., 1998). To date, the only metazoan histone fusion has been with a core histone, H2B. Although the core histones are responsible for forming the basis of the nucleosome, the condensation of the chromatin fibre is dictated by the actions of the linker histones. These proteins are responsible for folding and maintaining a stable higher order fibre (Allan et al., 1981); the attachment of another molecule, GFP, on to a linker histone may be expected to influence these functions. To minimise the likelihood of an alteration in the function of H5, we generated a fusion onto the N-terminus of H5 as this domain is not strictly required for the formation of the higher order chromatin fibre (Allan et al., 1986).

The coding sequence of the histone H5 gene was PCR amplified using the following gene-specific primers which contain BspEI and XhoI restriction sites (underlined). The forward primer also contains a flexible glycine linker.

H5-GFP.for: 5'-ATCCGTCCGGAGGTGGCGGGATGACGGAGAGCCTGGTC-3'
H5A.bak: 5'-GGCCGCTCGAGTTACTTCAGCTCACTTCTTCTTGGGCGATTT-3'

After digestion, the H5 fragment was ligated in-frame into pEGFP-C1 (Clontech), under the control of the strong immediate early promoter of cytomegalovirus. Individual clones were isolated and inserts were confirmed by sequencing, resulting in the protein EGFP-H5 (Figure 5.12). For ES cell expression studies, the GFP-H5 fusion construct was subcloned from pEGFP-H5 using NheI and EcoRI. The overhanging ends were filled using Klenow DNA polymerase and were ligated to BstXI/Blunt adapters and cloned into pHPCAGGS-BstXI to generate pHPCAGGS-GFPH5.

Before transfecting the GFP-H5 fusion protein into ES cells it was necessary to demonstrate that the fusion protein was functional and had the same properties as wild-type H5. This was undertaken in NIH3T3 cells and Cos7 cells.

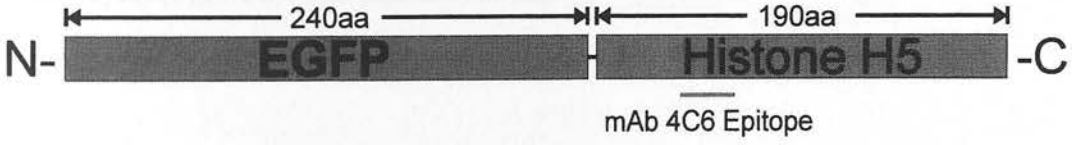


Figure 5.12 A schematic showing the GFP-Histone H5 fusion protein. PCR amplified histone H5 was ligated in frame with EGFP. The resulting 430 amino acid (aa) protein is shown. The histone H5 domains comprise an N-terminal extended domain (1-25aa), a compact globular central domain (26-100aa) and a long C-terminal tail (101-190aa). The epitope of the H5-specific monoclonal antibody, 4C6, is marked at the N-terminal end of the globular domain.

Expression of GFP-H5 fusion protein in eukaryotic cells

Two independent pEGFP-H5 clones were transiently transfected into NIH3T3 cells using liposome-mediated transfer. After 24 hours, examination of the cells by fluorescence microscopy showed dense green fluorescence localised to the nucleus (Figure 5.13C and 5.13D). In contrast, the control vector, pEGFP-C1, showed diffuse staining throughout the cells (Figure 5.13A, 5.13B; Chalfie et al., 1994). At higher magnification (Figure 5.13E and 5.13F), bright punctate foci are visible, indicating localised concentrations of GFP-H5 in the nucleus. These are either more compacted chromatin regions (Manuelidis, 1990), or regions of the chromatin which have a preference for binding GFP-H5. Most transfected cells expressed GFP-H5 at a similar level suggesting that in the short period of transient transfection, the fusion protein did not have a deleterious effect on the cells.

To analyse the GFP-H5 fusion protein at different expression levels I transfected linearised and closed circular pEGFP-H5 plasmid into three different cell lines. The cell lines employed were NIH3T3 cells, human 293 cells which constitutively express transcription factors for high level CMV promoter activity, and Cos7 cells which allow episomal replication of the pEGFP-H5 construct which has an SV40 ori. Cos7 and 293 cells would therefore be expected to give higher levels of GFP-H5 expression than NIH3T3 cells, although in all cell types used the expression level of this fusion protein is substantially less than the normal level of H5 in chicken erythrocytes (Figure 5.1 and data not shown).

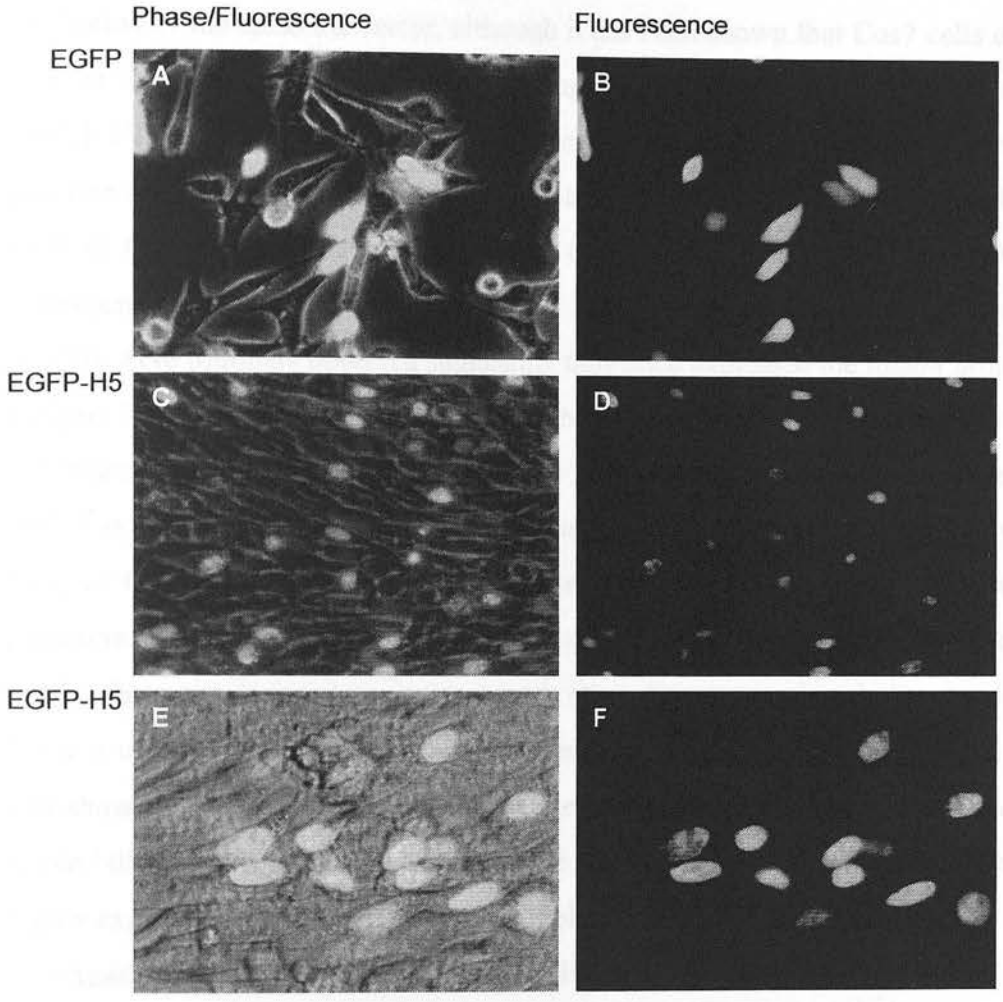


Figure 5.13 GFP-H5 is nuclear localised in NIH3T3 cells. NIH3T3 cells were transiently transfected with pEGFP-C1 or pEGFP-H5 for 24 hrs. (A, B) EGFP fluorescence diffusely stains the interior of the cells whilst EGFP-H5 fluorescence is localised to the nucleus (C, D). (E, F) At higher magnification, the EGFP-H5 fluorescence is not evenly distributed throughout the nucleus.

As seen for NIH3T3 cells, both 293 and Cos7 cells transiently expressed the GFP-H5 fusion protein and localised it to the nucleus as determined by fluorescence microscopy. However, pEGFP-H5 transfected Cos7 and 293 cells did not form stably expressing GFP-H5 clones after selection with G418. Presumably, this is due to high level expression of the fusion protein (or, for Cos7 cells, a toxic effect from runaway replication of the episomal vector, although it has been shown that Cos7 cells do not necessarily over-replicate plasmids carrying the SV40 origin (Chittenden et al., 1991)). NIH3T3 cells, transfected in a similar manner to the Cos7 and 293 cells, did generate stable cell lines and a mixed population of these clones was expanded. The level of GFP-H5 expression in the stably transfected NIH3T3 cells was low and heterogeneous although a large number of dead and dying cells which fluoresced brightly were routinely observed suggesting they once expressed the fusion protein at a higher level. The observed growth rate of NIH3T3 cells stably expressing the GFP-H5 fusion protein appeared to be lower than for the untransfected cells, in agreement with Sun et al. (1990c). Taken together these results suggest that a relatively low level of GFP-H5 expression can be tolerated by all the cells tested, albeit with a reduction in growth rate. However, high level expression of GFP-H5 leads to cell death. These results are in agreement with other data concerning the expression of H5 in a variety of cell types (Figure 5.7; Bergman et al., 1988; Aubert et al., 1991), and show that GFP-H5 does not inhibit the toxic effect of H5 expression. As the stable NIH3T3 cell lines only expressed the fusion protein at a low level I used the higher expressing transient transfections of pEGFP-H5 for further experiments.

Biochemical characterisation of the GFP-H5 fusion protein

To establish whether the GFP-H5 fusion had similar biochemical characteristics to native histone H5, Cos7 cells were transiently transfected with plasmid DNA carrying the fusion gene for 24 or 72 hrs. By 72 hrs the majority of cells had died, although surviving transfected cells continued to express GFP-H5 at a high level (as determined by fluorescence microscopy). Cells harvested at 24 hrs were extracted with 5% perchloric acid (PCA) (Section 2.10.1), a standard procedure for the selective isolation of linker histones (Van Holde, 1988), which normally partition to the soluble phase. Total proteins from cells harvested at 72 hrs were analysed directly

by SDS-PAGE as were the soluble and insoluble fractions from the PCA-extracted cells. The GFP-H5 fusion protein was detected by western blotting with the H5-specific monoclonal antibody, 4C6 (see Figure 5.12 for epitope). The results demonstrate that the GFP-H5 fusion protein, unlike native H5, is insoluble in 5% PCA (Figure 5.14, lane 4). Analysis of GFP-H5 by SDS-PAGE (Figure 5.14, lane 2) from a total cell extract shows the protein is approximately 48kDa as expected.

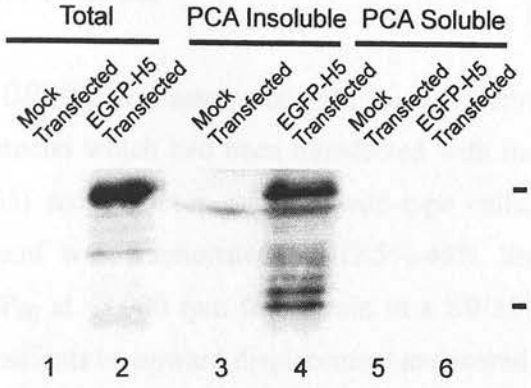


Figure 5.14 Recovery and PCA extraction of GFP-H5 from COS7 cells. COS7 cells were either mock transfected (no plasmid) or transfected with pEGFP-H5. Proteins from cells were fractionated by 15% SDS-PAGE, blotted and probed with the H5-specific monoclonal antibody, 4C6. Lanes 1 and 2, cells were harvested at 24 hrs post-transfection and total proteins analysed. Expressed GFP-H5 protein is clearly visible (top arrow). Proteins from cells harvested at 72 hrs were PCA extracted and the soluble (Lanes 5 and 6) and insoluble (Lanes 3 and 4) phases were analysed. The GFP-H5 protein is recovered in the PCA-insoluble fraction. The position of wild-type H5 protein is marked for reference (lower arrow).

Functional characterisation of the GFP-H5 fusion protein

The most rigorous test to show that GFP-H5 functions correctly would be to demonstrate that it is able to protect a 168 base pair chromatosome in an *in vitro* assay (Allan et al., 1986). Normally this would involve reconstituting the purified linker histone onto stripped (linker histone-depleted) chromatin and then digesting with micrococcal nuclease. However, the expression level of the GFP-H5 fusion protein in eukaryotic cells was too low to provide sufficient material and to date it has not been possible to express full-length histone H5 in bacteria due to features inherent in the linker histone C-terminal tail (Gerchman et al., 1994). As I was unable to express recombinant GFP-H5 in sufficient amounts for an *in vitro* assay my other studies on the fusion protein were carried out by transfecting Cos7 cells with the GFP-H5 expression construct.

To demonstrate that GFP-H5 was associated with the chromatin fibre, chromatin was isolated from Cos7 nuclei which had been transfected with the GFP-H5 expression construct (pEGFP-H5) for 24 hours or from wild-type cells. The chromatin was released overnight and was fractionated by 12.5%-45% linear sucrose gradient sedimentation in TEP_{80} at 48,000 rpm for 90 min in a SW55 rotor. Fractions were collected from the gradients by upward displacement and stored at 4°C. To determine the relative level of GFP-H5 bound to the chromatin, the fluorescence of each fraction was measured using a fluorimeter at 488nm excitation and 507nm emission (Figure 5.15). In addition, proteins were precipitated from the gradient fractions and analysed by SDS-PAGE to confirm that the GFP-H5 fusion was present (data not shown). As expected, the endogenous fluorescence of the untransfected cells is very low demonstrating that the signal seen in the GFP-H5 expressing cells is not attributable to other cellular components that may have contaminated the preparation and bound to chromatin (Figure 5.15A). For the GFP-H5 transfected cells, fluorescence is broadly correlated with the amount of chromatin (Figure 5.15B, Fractions 4-9), suggesting that the GFP-H5 fusion protein is incorporated into the chromatin fibre in an appropriate, stoichiometric fashion. The large peak at the top of the gradient (Figure 5.15B, Fractions 2 and 3) is probably due to RNA or free DNA fragments present in the preparation which will competitively bind GFP-H5, and some protein degradation that would release free GFP.

GFP-H5 binds to nucleosomal particles

To further characterize the interaction of GFP-H5 with the chromatin fibre the association of the fusion protein with individual nucleosomes was analysed. Soluble polynucleosomes were prepared from Cos7 cells transfected with either pEGFP-C1 or pEGFP-H5. Soluble Cos7 chromatin was dialysed overnight against TEP_{80} and then further digested with micrococcal nuclease to prepare short oligo-nucleosomes (Figure 5.16A). After addition of ficoll to 3.5% (w/v), the chromatin fragments were fractionated on a 5% polyacrylamide gel (Section 2.8.2) in TBE at 4°C. The gel was analysed for GFP fluorescence using a scanner equipped with a 479nm laser and a 520nm bandpass filter (Figure 5.16B). The gel was stained with ethidium bromide and further scanned using a 479nm laser and a 580nm bandpass filter (data not shown). Control cells (Figure 5.16B, Lane 1) transfected with pEGFP-C1 did not display fluorescence indicating that native GFP does not bind to the chromatin fibre (although, only a very small fraction of the GFP protein would be nuclear localised; Figure 5.13A, 5.13B). In contrast, a clear nucleosomal banding pattern is visible for the GFP-H5 containing nucleosomes (Figure 5.16B, Lane 2). In addition to the mono-nucleosome band, dimer, trimer and tetramer bands are clearly visible indicating that the fusion protein is associated with individual nucleosomes at approximately stoichiometric levels (compare Figure 5.16A lane 2 to Figure 5.16B lane 2).

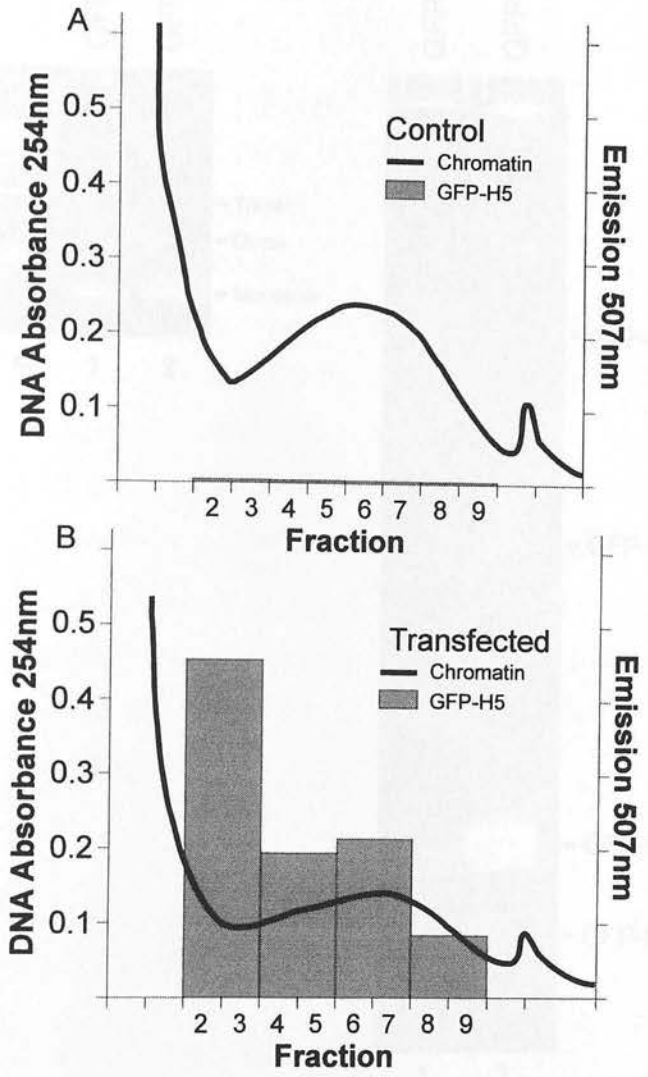


Figure 5.15 GFP-H5 is associated with the chromatin fibre. COS7 cells were either mock transfected (no plasmid) or transfected with pEGFP-H5 for 24 hrs. Soluble chromatin was recovered and sedimented on a 12.5-45% linear sucrose gradient. The GFP fluorescence in each fraction was measured. (A) Mock transfected cells. (B) Cells expressing GFP-H5.

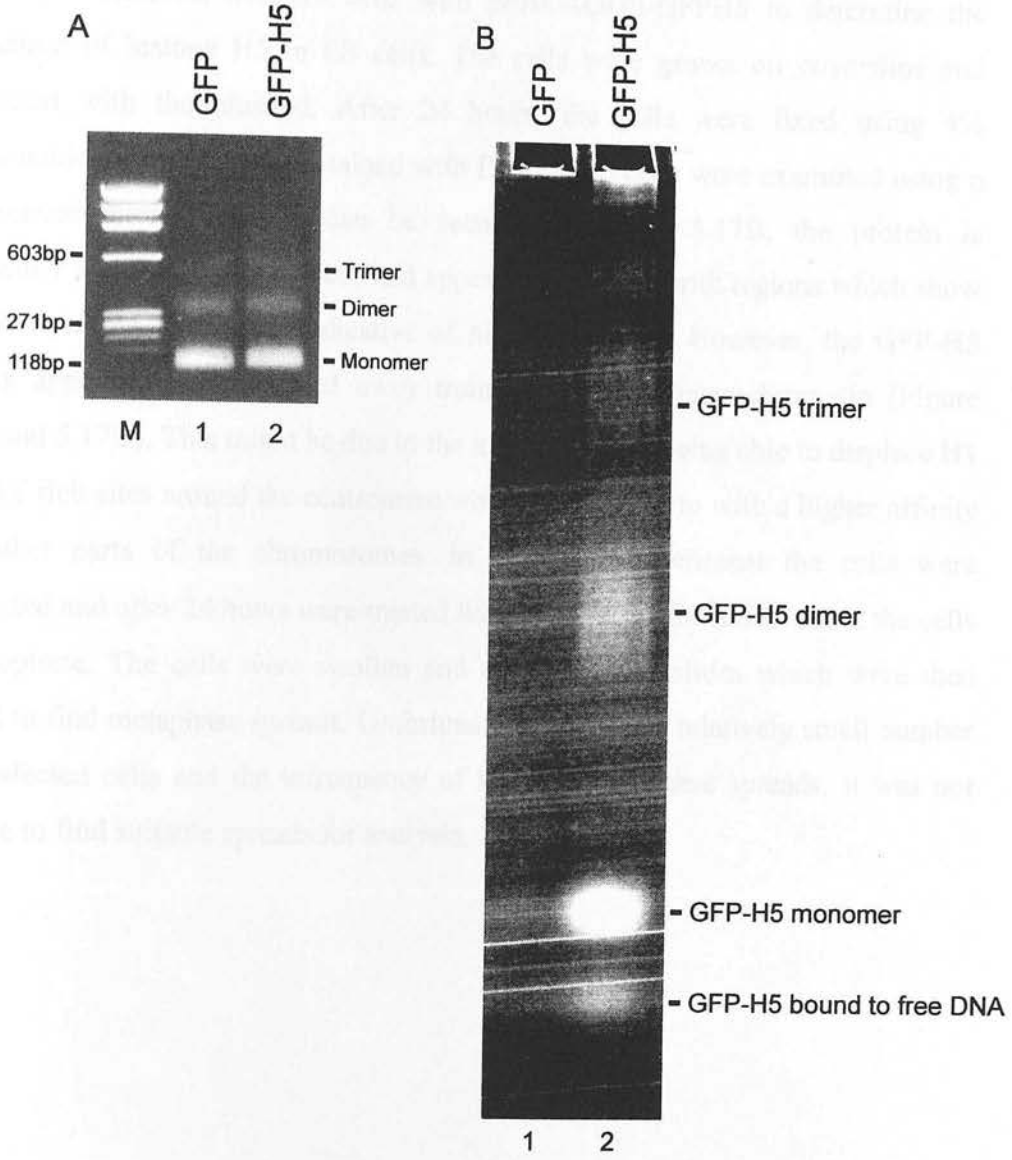


Figure 5.16 GFP-H5 is associated with individual nucleosomes. COS7 cells were either transfected with pEGFP-C1 or pEGFP-H5. Soluble chromatin was recovered and digested with micrococcal nuclease before analysing on a nucleoprotein gel. GFP fluorescence was detected by laser-scanning at 479nm. (A) 1% agarose gel analysis of DNA isolated from chromatin after micrococcal nuclease digestion. Markers (M) are ØX174 HaeIII. (B) Chromatin analysed on a native 5% polyacrylamide gel. Lane 1, chromatin from cells transfected with pEGFP-C1, Lane 2, chromatin from cells transfected with pEGFP-H5.

GFP-H5 localisation in ES cells

Having demonstrated that the GFP-H5 fusion protein functions similarly to native H5, I proceeded to express this protein in ES cells. As a preliminary experiment, I transiently transfected MG1.19 cells with pPHCAGGS-GFPH5 to determine the localisation of histone H5 in ES cells. The cells were grown on coverslips and transfected with the plasmid. After 24 hours, the cells were fixed using 4% paraformaldehyde and counter stained with DAPI. The cells were examined using a fluorescence microscope. As can be seen from Figure 5.17B, the protein is specifically localised to the nucleus and appears to coincide with regions which show more intense DAPI staining indicative of heterochromatin. However, the GFP-H5 staining appears to be localised away from centromeric heterochromatin (Figure 5.17C and 5.17D). This might be due to the inability of H5 being able to displace H1 from AT rich sites around the centromere which it will bind to with a higher affinity than other parts of the chromosomes. In a similar experiment the cells were transfected and after 24 hours were treated with 0.1μg/ml colcemid to arrest the cells in metaphase. The cells were swollen and cytospun onto slides which were then studied to find metaphase spreads. Unfortunately, due to the relatively small number of transfected cells and the infrequency of forming metaphase spreads, it was not possible to find suitable spreads for analysis.

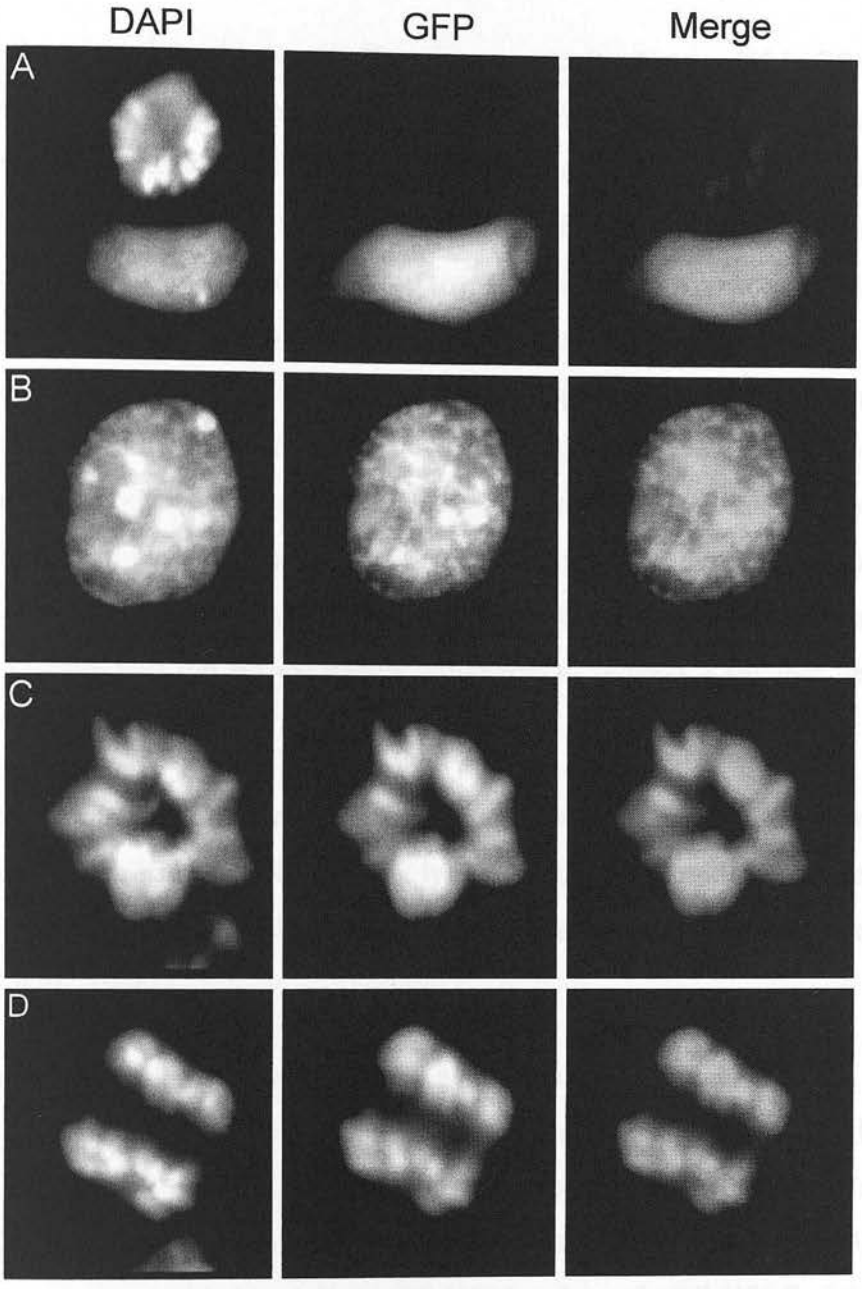


Figure 5.17 GFP-H5 is localised away from centromeric heterochromatin in ES cells. MG1.19 ES cells were grown on coverslips and were transfected with pEGFP-C1 or pPHCAGGS-GFPH5. After 24 hours the cells were fixed and counter-stained with DAPI. (A) GFP fluorescence is distributed throughout the cell. (B) GFP-H5 tends to be localised to regions which show intense DAPI staining indicative of heterochromatin, however the images of cells in a prometaphase rosette (C) and during anaphase (D) suggest that GFP-H5 is excluded from centromeric heterochromatin.

5.5 Histone H5 DNA binding mutants are lethal to ES cells

Although there is a clear phenotypic effect from expressing histone H5 in ES cells it is important to provide suitable negative controls for these experiments. Histone H5 is not normally found in ES cells and, in conjunction with it being an extremely positively charged protein, it might have non-specific effects causing an altered phenotype. Previously, a related linker histone, H1°, has been over-expressed in ES cells to make a transgenic mouse with no overt phenotypic effect (Tonjes et al., 1997), making it unusable as a control in these experiments. The most suitable negative control would be an H5 mutant which is unable to interact properly with the chromatin fibre. Goytisolo et al (1996) have mutated the putative DNA binding residues in the globular domain of histone H5 and showed that the mutant protein is unable to interact correctly with the chromatin fibre. I therefore cloned these mutated globular domains into the H5 expression constructs to provide a suitable negative control for my experiments.

The globular domain of H5, in my expression constructs, was replaced with a mutant H5 fragment which had the DNA-binding lysine and arginine residues mutated to a mixture of glutamines and alanines (glu+ala) or only alanines (all ala) (Figure 5.18; Goytisolo et al., 1996).

The mutant globular domains were made in a previous study (Goytisolo et al., 1996; Goytisolo and Ramakrishnan, unpublished) and were subcloned as EaeI/MscI fragments into pBS-H5for by a three-way ligation to generate pBS-H5(glu+ala) and pBS-H5(all ala). These constructs were confirmed by sequencing (Figure 5.19). The mutant H5 genes were subcloned from pBS-H5(glu+ala) and pBS-H5(all ala) using EcoRI and were ligated to EcoRI/BstXI adapters and cloned into pHPCAGGS-BstXI to generate pHPCAGGS-H5(glu+ala) and pHPCAGGS-H5(all ala). In addition, H5 was subcloned from pBS-H5for at the EcoRI site and ligated to EcoRI/BstXI adapters (Section 2.4.5) and cloned into pPHCAGGS-BstXI to generate pPHCAGGS-H5Bst.

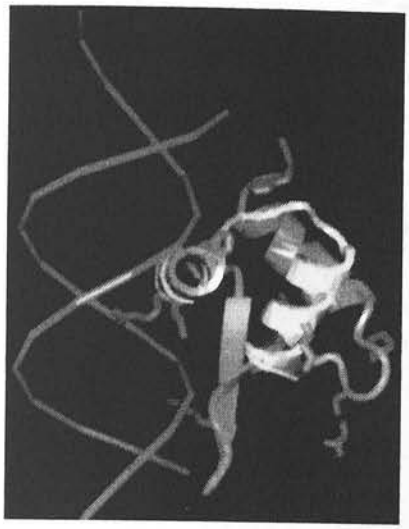
A

Mutations

glu + ala:K40E, R42E, K52A, K69A, R73A, K85A, R94A

all ala:K40A, R42A, K52A, K69A, R73A, K85A, R94A

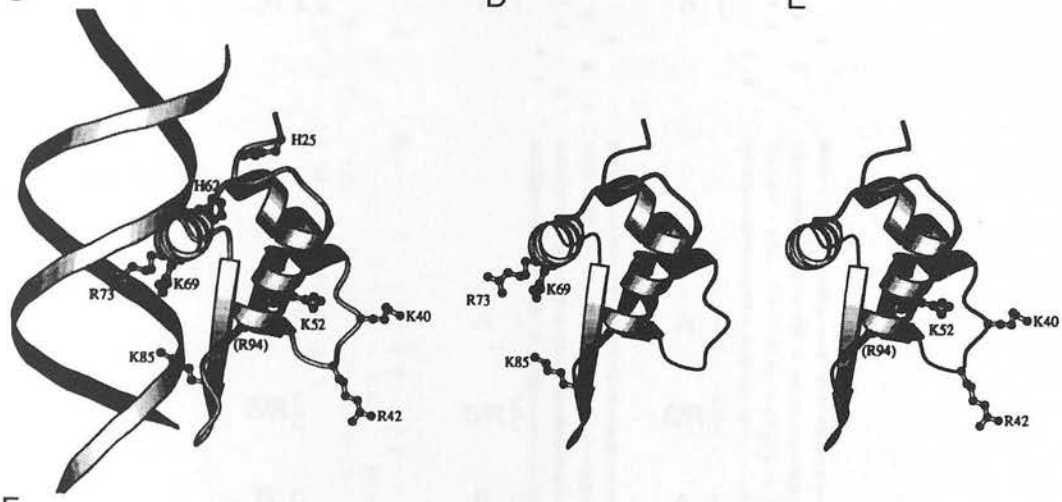
B



C

D

E



F

Figure 5.18 Native and mutant histone H5 globular domain DNA-binding sites. Histone H5 globular domain has two DNA-binding sites composed of positively charged amino-acids. (A, F) These key binding sites were altered to inhibit the correct binding of H5 to the nucleosomal linker DNA (B-E). (Pictures taken from Goytisolo et al., 1996).

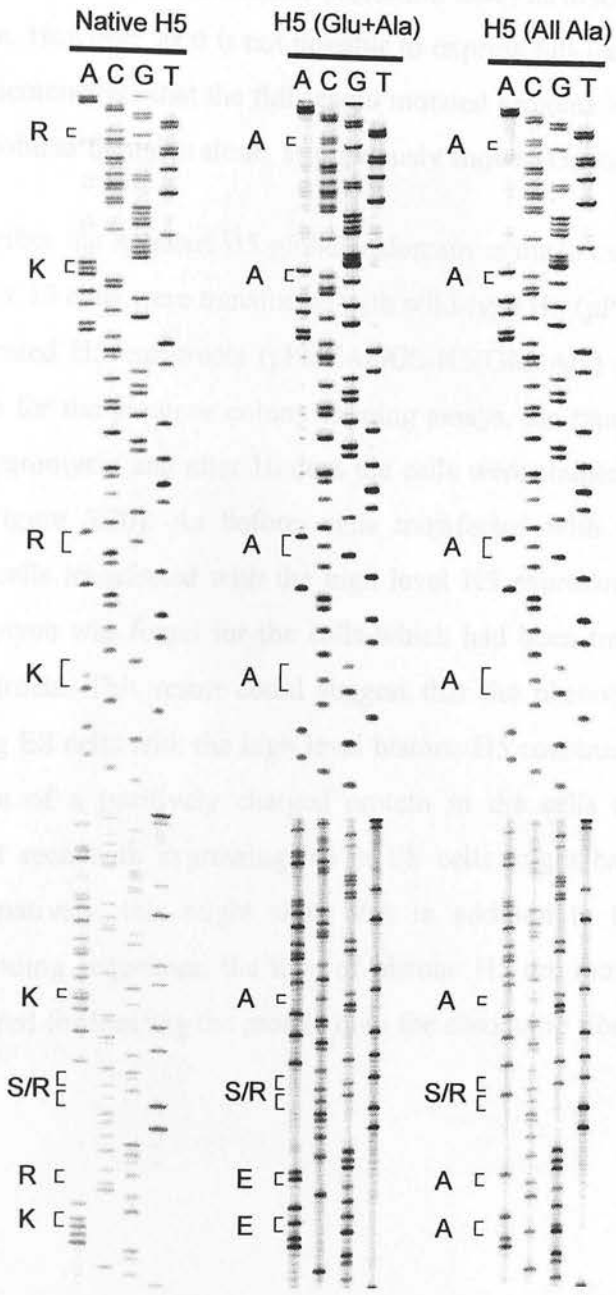


Figure 5.19 Sequence analysis of H5 DNA-binding domain mutants. To confirm the full-length mutant H5 proteins carried the necessary mutations the genes were analysed by DNA sequencing. Silent mutations generated during the construction of the H5 DNA-binding mutants are marked as S/R.

To conclusively demonstrate that the mutated H5 constructs had the expected phenotype of being unable to interact with the chromatin fibre it would be necessary to use purified protein in a chromatosome protection assay as discussed for the GFP-H5 fusion protein. However, as it is not possible to express full-length H5 in bacteria I was unable to demonstrate that the full-length mutated proteins had a similar effect to the mutated globular domains alone, as previously shown (Goytisolo et al., 1996).

To establish whether the mutated H5 globular domain mutants exerted an effect on the ES cells, MG1.19 cells were transfected with wild-type H5 (pPHCAGGS-H5Bst) and the two mutated H5 constructs (pPHCAGGS-H5(Glu+Ala) and pPHCAGGS-H5(All Ala)). As for the previous colony forming assays, the transfected cells were selected using hygromycin and after 10 days the cells were stained with Leishman's fix and stain (Figure 5.20). As before, cells transfected with the empty vector survived whilst cells transfected with the high level H5 expressing construct died. This same phenotype was found for the cells which had been transfected with the mutant H5 constructs. This result could suggest that the phenotypic effect found when transfecting ES cells with the high level histone H5 construct was merely due to the expression of a positively charged protein in the cells and therefore the phenotypic effect seen with expressing H5 in ES cells might have been partially artifactual. Alternatively, this might show that in addition to the core globular domain DNA binding sequences, the tails of histone H5 are more important than currently considered for locating the protein onto the chromatin fibre and exerting its effect.

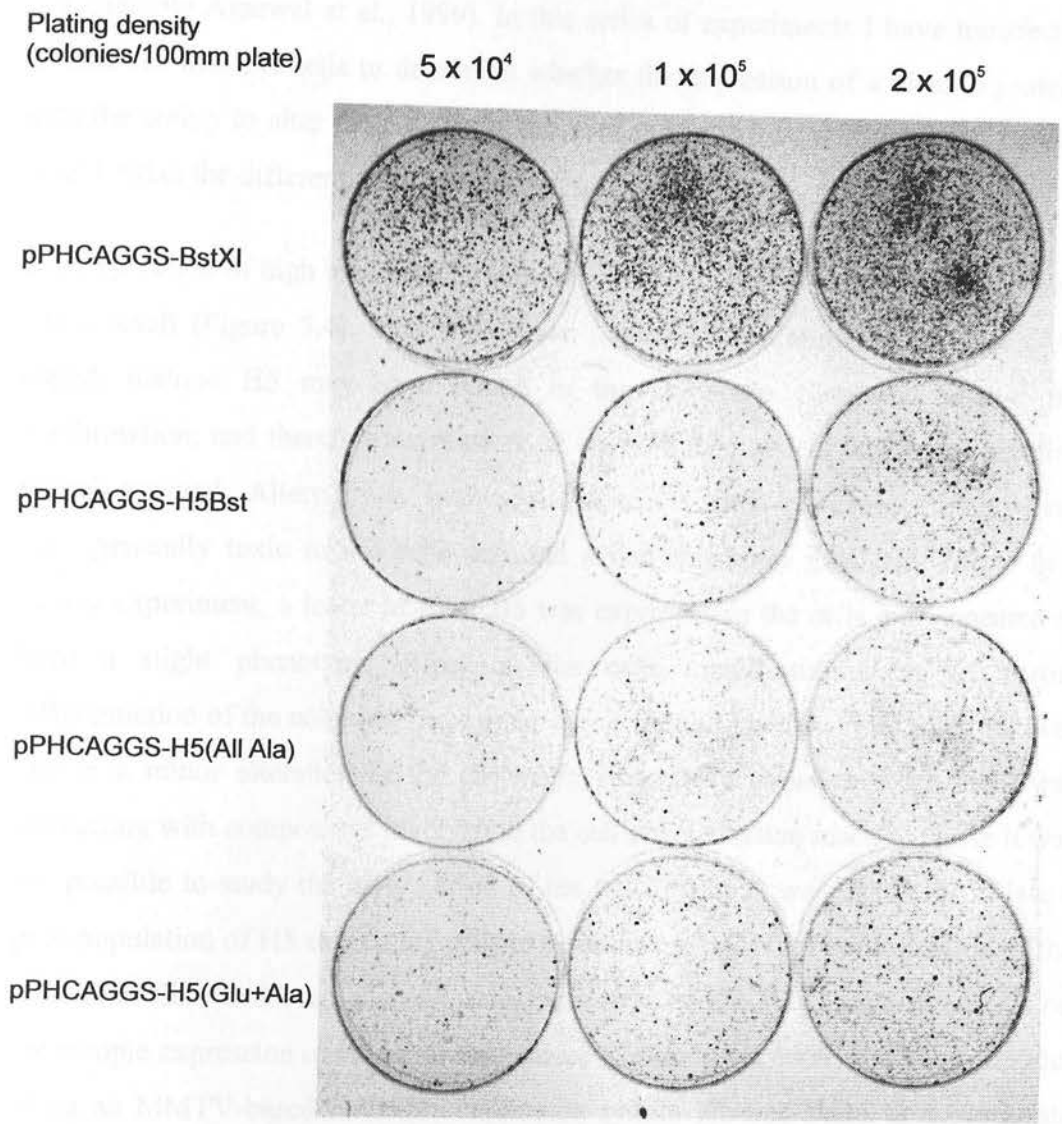


Figure 5.20 Histone H5 DNA-binding mutants are lethal to ES cells. MG1.19 cells were transfected with control vector (no insert), pPHCAGGS-H5Bst expressing H5 and the two mutant constructs. The cells were plated at various densities and after 24 hours hygromycin was added to the medium to select for transfected cells. After 10 days the plates were stained with Leishman's fix and stain to visualise surviving colonies.

5.6 Discussion

During the process of cell differentiation a number of key long-range chromatin changes are likely to occur to facilitate the co-ordinated expression of required genes (for example Agarwal et al., 1999). In this series of experiments I have transfected histone H5 into ES cells to determine whether the expression of a histone protein, with the ability to alter the chromatin structure at a global level (Sun et al., 1990c), would affect the differentiation of these cells.

The expression of high levels of H5 during the early period of the transfection gave a lethal result (Figure 5.4). This could have occurred for a number of reasons. As hoped, histone H5 may have bound to the chromatin fibre and altered the conformation, and therefore expression, of a wide range of loci which are required for cell survival. Alternatively, because the protein is positively charged it may have been generally toxic to the cells and had a non-chromatin mediated effect. In a further experiment, a lower level of H5 was expressed in the cells and appeared to have a slight phenotypic effect on the cells, manifesting itself as partial differentiation of the colonies. Once again it is difficult to conclude whether this was due to a minor alteration in the chromatin structure or whether the protein was interacting with components involved in the cell differentiation machinery. As it was not possible to study the early period of the transfection, I was unable to isolate a pure population of H5 expressing cells to determine whether the conformation of the chromatin fibre had been altered by the expression of H5. Previous experiments on the ectopic expression of H5 in rat sarcoma cells (Sun et al., 1989; Sun et al., 1990c) using an MMTV-based inducible expression system allowed them to demonstrate that the conformation of the chromatin fibre was altered by the expression of H5.

In my experiments, low level stable H5 expression was detected in the few clones which survived the transfection (Figure 5.6) suggesting that some cells were able to compensate for an alteration in the chromatin structure or they were able to down-regulate the H5 expression early in the transfection before the level of H5 became toxic. Histone H5 could be transfected into NIH3T3 cells suggesting that ES cells may be more sensitive to the effects of H5. This could be due to NIH3T3 cells having

a more robust chromatin architecture which is less prone to H5-mediated gene inactivation or because rapidly proliferating ES cells which are finely balanced between stem cell renewal and subsequent differentiation are more prone to H5 mediated non-chromosomal interference to other cellular pathways.

To ensure the phenotypic effects observed for the expression of H5 in ES cells were due to an alteration in the chromatin fibre, I felt it was important to examine cells early in the transfection. The two approaches planned were to express histone H5 under the control of a tetracycline-based expression system and to express a fluorescently tagged H5 molecule which would allow the isolation of the transfected population of cells. The tetracycline-based expression system, although constructed, was never fully tested. Very preliminary studies did not give any H5 expression upon tetracycline withdrawal but this is likely to be due to the way the experiment was carried out rather than because the system was not functioning correctly (Niwa et al., 1998). The use of the tetracycline-based expression system is now being superseded by the Cre-lox (Sauer and Henderson, 1990; Smith et al., 1995) expression systems. In one of these approaches there is a stop signal flanked by loxP sites inserted between the promoter and the gene such that transcription is prematurely terminated (Mao et al., 1999). Activation of Cre in the cells by a regulated Cre fusion protein (for example, Cre-tam, which is activated by tamoxifen (Littlewood et al., 1995; Vasioukhin et al., 1999) causes the loxP-flanked stop signal to be excised and the gene of interest to be expressed. The advantage of a Cre-lox controlled expression system compared to the tetracycline expression system which I used is that it would be relatively easy to express a variety of different constructs in the ES cells without having to repeatedly target the tetracycline cell line. Conversely, the disadvantage of this system is that there is no regulation of the level of gene expression (except by the promoter), so for some applications where it might be more informative to give a burst of gene expression it would be unsuitable.

The second approach for selecting a population of cells which express H5 is by transfecting the cells with a GFP-H5 fusion protein. Characterisation of the fusion protein has demonstrated that it functions as H5 and would therefore be suitable for

these experiments (Section 5.4). The approach I had planned was to transfect cells with the fusion protein and to visualise these cells shortly after transfection to determine whether there was any alteration in cell phenotype. Also, expressing a GFP fusion protein would have allowed the transfected cells to be isolated from the total population of cells for further, more detailed study.

One of the observations made for ES cells, NIH3T3 and Cos7 cells transfected with GFP-H5 is that some cells within the cell population expressed high levels of the fusion protein and appeared to display a cellular phenotype reminiscent of cells undergoing apoptosis, including a blebbing membrane, condensed nucleus, and many intracellular vesicles. Although this avenue was never pursued, it would be interesting to determine whether the expression of H5 was inducing some-kind of apoptotic response in the ES cells, as opposed to a necrotic response. This would not be unprecedented as previous studies (Jayaraman et al., 1998) have shown that HMG-1 interacts with p53 showing there is a direct link between components of the apoptosis detecting machinery and elements of the chromatin fibre. Also, as an alteration in the conformation of the chromatin fibre would be expected to be detrimental to cells, a detection mechanism for identifying this cellular change would not be unlikely. ES cells have been shown to be especially sensitive to DNA damage and the induction of apoptosis (Van Sloun et al., 1999; Corbet et al., 1999) and that this decreases as stems cells differentiate (Sabapathy et al., 1997). If H5 were causing an apoptotic response, resulting in the death of the cells, then the use of a $p53^{-/-}$ ES cell line may be expected to minimise these effects, although contradictory results suggest ES cells can undergo p53-independent apoptosis in response to DNA damage (Aladjem et al., 1998). Repeating these experiments with a $p53^{-/-}$ ES cell line would not be trivial as the episomal expression system used in these experiments would be difficult and time-consuming to make in another cell line. Likewise, the generation of a homozygous p53 knockout in the MG1.19 cell line would also be very time-consuming and difficult, especially as these cells already carry a number of the commonly employed resistance markers thus limiting the number of further markers which could be used. An alternative approach would have been to pick the surviving

clones from the H5 transfection (Figure 5.6) and, assuming these cells to be resistant to the toxic effect of H5, re-transfect these cells to see if they now behave differently from the original transfection.

Expression of a fluorescent H5 fusion protein in cells was intended to allow the nuclear localisation of H5 to be observed and it may also prove to be useful for future studies on the expression of linker histones during development. My studies on the expression of H5 in interphase ES cells showed that the protein was localised to specific regions of the chromosomes (Figure 5.17). It was not possible to find a metaphase spread from a transfected cell as the combined frequency of transfection and formation of a spread were too low, so I was unable to accurately determine the localisation of H5. One current view on the localisation of linker histones (Crane-Robinson, 1999) is that H1 subtypes may have a more localised distribution whereas linker histones like H1° and H5 are likely to be more globally distributed. Treatment of cells with 5-azacytidine has been shown to decondense metaphase chromosomes (Joseph et al., 1989) allowing the localisation of chromatin-associated proteins to be determined at higher resolution. If an alternative expression system was used which allowed a larger proportion of the cells to be expressing the fusion protein, it would be possible to study the localisation of H5 to determine its ectopic expression pattern in detail.

To date it has been shown that patterns of linker histones expressed or present in a cell alter during development. These patterns of protein expression have been characterised extensively during the early years of histone protein research (Van Holde, 1988) but to date, detailed studies have not been undertaken to determine the expression pattern of linker histones in developing organisms. One use of a linker histone fusion protein would be to study the distribution of a specific linker histone during the development of a suitable organism, such as zebrafish. Currently, zebrafish transgenics (Lin et al., 1994) are in their infancy but it is now possible to generate stable zebrafish transgenics at higher frequency (Gaiano et al., 1996; Linney et al., 1999) and it has been shown that GFP fluorescence can be monitored during their development (Amsterdam et al., 1995; Higashijima et al., 1997). A GFP linker

histone fusion protein could be transfected into zebrafish with appropriate control sequences to allow the expression of linker histones to be explored in real-time during the development of specific systems.

My experiments for the alteration of ES cell chromatin structure by the expression of histone H5 were never taken to fruition as an H5 DNA-binding domain mutant appeared to have the same phenotypic effect as native H5 (Figure 5.20). As mentioned, this is possibly because histone H5 is causing non-chromatin-mediated interactions which are incompatible with the survival of ES cells. Alternatively, the interactions of H5 with the chromatin fibre are not fully understood and it might be that these mutants are still able to interact with the chromatin fibre and mediate an effect. Many of the studies to date on the interaction of H5 with the chromatin fibre have been carried out *in vitro* under controlled situations (Allan et al., 1981; Thomas and Wilson, 1986) and it maybe that *in vivo* there are more factors which have not been taken into account (Ali and Singh, 1987). If the lethal effects of H5 are caused by triggering an apoptotic response it is quite likely that the DNA binding mutants would also induce it. Although the binding mutants are not able to protect a chromatosome they do bind to DNA (Goytisolo et al., 1996), and depending on the cellular apoptosis detection system may also trigger a response.

In light of my current results I feel that the expression of H5 in ES cells as a route to modulate the chromatin structure may be prone to influence by other factors, and without characterising these effects it would be difficult to draw any conclusions from my studies. If these experiments were repeated then another approach should be used which would allow stable cell lines which carried the H5 gene to be generated and in which the expression of the transgene could be readily activated. One example would be to integrate a Cre-tam expression construct in the MG1.19 ES cells and to insert suitable stop sequences flanked by loxP into my expression vectors. It would thereby be possible to express H5 co-ordinately in a population of cells. This would allow a detailed analysis of the effects of H5 and, most importantly, would allow the chromatin to be isolated to examine whether there was a change in its conformation both for the wild-type and mutant H5 proteins. By integrating this approach with the

GFP fusion construct, it would allow a direct correlation to be drawn between the expression of H5 in these cells and any possible change in the level of apoptosis. Previous data have suggested that it is possible to estimate the extent of apoptosis within a population of cells purely by FACS, depending on the incorporation of Hoechst 33342 to estimate chromatin condensation and propidium iodide staining to assess membrane damage (Belloc et al., 1994). However, the use of a fluorescent-based TDT-mediated dUTP-biotin nick end-labelling approach or annexin V analysis is more accurate and could be correlated to the level of H5 expression in individual cells by FACS (O'Brien et al., 1997). Thus a correlation could be drawn between the extent of apoptosis, the level of H5 expression and the conformation of the chromatin fibre.

6. Conclusions

The conformation of the higher order chromatin fibre is important for regulating the expression of genes. For many years the fibre has been studied, but still its structure remains poorly understood (Ramakrishnan, 1997). In this thesis, I have studied the conformation of the fibre in two situations. Firstly, I have shown that the conformation of the higher order chromatin fibre does not alter when cells are differentiated, in contrast to the widely held notion that it does, and secondly I have demonstrated that the chromatin associated with centromeres has a special conformation compared to bulk chromatin, an observation which has not been previously made.

During my studies on stem cell and differentiated cell higher order chromatin I have been unable to identify any differences between their bulk fibre conformations. This result is important because it demonstrates that the chromatin changes that happen during lineage determination occur at a local as opposed to a bulk level and are mediated by factors other than the conformation of the higher order fibre. In agreement with previous data from our laboratory which studied the conformation of the chicken β-globin gene (Caplan et al., 1987), I was unable to find a difference in the chromatin structure of the developmentally important Oct-4 gene and bulk chromatin in ES cells. This suggests that even at the level of specific genes there is no change in the higher order chromatin conformation. Many recent studies have shown that gene expression is regulated by large chromatin remodelling complexes (Workman and Kingston, 1998) which facilitate nucleosome re-organisation and histone acetylation, but these local alterations probably do not affect the conformation of the higher order fibre. My studies show there is no alteration in the conformation of the higher order fibre at both a bulk level and the level of specific genes in undifferentiated ES cells and differentiated daughter cells even though there is likely to be substantial re-organisation of local genes by chromatin remodelling activities. This suggests it is unlikely that either large nucleosome remodelling complexes or acetylation are influencing the conformation of the higher order fibre.

My results have shown that although there is not a change in the conformation of the bulk chromatin fibre in stem cells and differentiated cells there is a difference in the nuclease sensitivities of these cell types. As mentioned previously (Section 3.6), it is difficult to interpret these results as the conformation of the higher order fibre is poorly understood and therefore it is not clear what specific elements of the fibre are being probed by nucleases.

An interpretation of these results is that although there is no alteration in the conformation of the higher order fibre there is likely to be an alteration in the stability or equilibrium of the fibre. Little data supports the idea that either nucleosome remodelling complexes or histone acetylation affect the conformation of the fibre, but some data does suggest that they are able to alter the stability of the system (Tse et al., 1998). This idea is reasonable, as it is clear that histones are modified by acetylation, methylation and phosphorylation at key sites in specific cellular circumstances suggesting that these alterations are important for regulating the conformation of the fibre. In addition, studies on specific genes such as the yeast HO locus have shown that for gene expression to occur it is necessary to first destabilise the nucleosomes to facilitate the access of acetylation complexes which then presumably stabilise this more open structure allowing transcription to proceed (Krebs et al., 1999; Cosma et al., 1999). In a nuclear environment the chromatin will be topologically constrained at both a domain level by the interaction of locus control regions with topoisomerase II and the nuclear matrix (Higgs, 1998) and at a more local level by the interaction of chromatin components like histone H2A and H2B by the cellular cytoskeleton (Goldberg et al., 1999). If histone acetylation and nucleosome remodelling only mediate an alteration in the chromatin structure within these torsional constraints the localised regions of altered chromatin structure will be maintained whilst the nuclei are intact. This can be probed by nuclease digestion, but once the constraints are removed as in the chromatin sedimentation analysis the influence of chromatin remodelling by acetylation, phosphorylation, methylation and nucleosome remodelling will be lost.

Previous work has shown that active and inactive genes have an equivalent higher order chromatin structure (Fisher and Felsenfeld, 1986). I have now shown that the higher order chromatin fibre of undifferentiated and differentiated cells has the same conformation. To further our understanding of the higher order fibre I have examined the conformation of the chromatin found at centromeric locations to establish whether it has a similar or different conformation to bulk chromatin. My results clearly show that the chromatin fibre found at centromeres is more compact, and that this compaction appears to be independent of DNA methylation, or histone acetylation. Interestingly the degree of compaction of the fibre appears to decrease away from the centromere suggesting that one role of the major satellite might be to buffer the condensed centromeric chromatin away from the bulk chromatin. There are number of determinants which could affect the compaction of centromeric chromatin; possible reasons include: positioning of nucleosomes on the underlying DNA sequence, replacement of H3 by CENP-A, or the presence of another centromere-associated protein. The next step in these experiments will be to investigate the molecular basis of the centromere conformation and find how this undoubtedly important alteration is mediated. The identification of functional neocentromeres which appear to have no sequence similarity to the native centromere suggest that the basis for the chromatin compaction is more likely to be a protein component rather than just the presence of a specific underlying DNA sequence. However, if it is found that the stable neocentromeres do have specific sequence motifs which allow the fibre to be packaged in a more condensed fashion, then it is possible that the positioning of the nucleosomes in a regular and ordered fashion may be responsible for maintaining the chromatin fibre in this conformation.

Modulation of the conformation of the chromatin fibre would be expected to alter the expression of many genes. Even though my data (Chapter 3) suggests that the conformation of the chromatin fibre was not altered during cell differentiation I thought it would be interesting to investigate the effect of modulating the chromatin structure of pluripotential embryonic stem cells. To study this, I expressed the linker histone H5 in ES cells. Few conclusions could be drawn from this work as it appeared that H5 DNA-binding mutants gave a similar phenotype to native H5. As

H5 is highly charged it is likely that both the native and mutant proteins have the ability to alter the chromatin fibre in some fashion. Therefore, rather than having a specific effect the results suggest that H5 was causing a non-specific, phenotypic change in the cells. With a low level expression of H5 it was possible to see a slight alteration in the differentiation of the ES cells potentially due to a direct effect from H5, but without being able to study the conformation of the chromatin fibre it was impossible to determine whether this was the case. Fluorescence microscopy of cells transfected with a GFP-H5 fusion protein suggested that the cells appeared to undergo an apoptosis-like event which, if followed up in future work, may indicate that the conformation of the higher order chromatin fibre is able to monitor its conformation and trigger an apoptosis-like pathway as some data suggest (Rogakou et al., 1999).

From my results on the expression of H5 in ES cells it may appear that the cellular alterations caused by H5 may have been too extreme. Previously, H1° has been expressed in ES cells, and H1° and H1c have been over-expressed in mouse fibroblasts with no apparent phenotypic effect, except for a small change in transcription (Tonjes et al., 1997; Brown et al., 1997b). Whilst H5 has been expressed in rat sarcoma cells and shown to alter the stability of the chromatin fibre (Sun et al., 1989; Sun et al., 1990c). This suggests that the approach of expressing H5 in these cells was valid, the disadvantage being that the expression system used was unsuitable for evaluating the results of the experiment. With hind-sight the first priority should have been the optimisation of an expression system which would have allowed the expression of H5 in a population of cells synchronously either using a (non-targeted) tetracycline regulated or Cre-lox expression system. This would have allowed the conformation of the chromatin fibre to be determined, to establish the effect native and mutant H5 was having on the chromatin fibre.

The results from my thesis suggest that chromatin found at centromeres has a unique conformation which is likely to be important for the cellular function of this chromosome component. Alterations in the chromatin structure between stem cells and differentiated cells were not found, although the alterations in the nuclease

sensitivity of the chromatin fibre suggests that there are differences, possibly in the dynamics, of the higher order fibres. Finally, my results suggest that modulation of the conformation of the chromatin fibre may alter the phenotype of the cells and that there are possibly cellular mechanisms for detecting alterations in the conformation of the chromatin fibre. Although, in future experiments which modulate the conformation of the chromatin fibre it will be important to ensure the effects seen are chromatin-specific. This study has provided suitable data which can be used for extending these projects allowing us to understand, in greater detail, how the higher order chromatin fibre is able to regulate cellular processes.

7.1 Plasmids

Number	Name	Description
16	pBR H5 2.4	H5 genomic fragment (Jerry Dodgeson)
20	pCEP4	Expression vector (Invitrogen)
23	pPHCAGGS-BstXI	CAG expression vector (Niwa et al., 1998; Burdon et al., 1999)
24	SuperKO	Tetracycline based targetting vector (Niwa et al., 1998)
25	pPHTK-H5	H5 expression construct
27	pPHCAGGS-H5	H5 expression construct
38	pPHTK-H5NX	H5 expression construct
39	pPHCAGGS-H5NX	H5 expression construct
47	R198	Mouse minor satellite (Kipling et al., 1994). 360 bp EcoRI/HindIII fragment.
54	pEGFP-C1	GFP fusion vector (Clontech)
77	pSAT	Mouse major satellite (Lewis et al., 1992). 240 bp EcoRI/BamHI fragment.
78	pPHTK-BstXI	TK expression vector
84	LVE(nx)soldr cDNA	Mouse B2 repeat (Chambers et al., 1997). 250 bp HindIII/NotI fragment.
91	pEGFP-H5	GFP-H5 expression construct
106	p82H	Human α-satellite (Mitchell et al., 1985). 2.2 kb EcoRI/HindIII fragment.
122	pPHCAGGS-H5Bst	H5 expression construct
123	pPHCAGGS-H5(Glu+Ala)	Mutant H5 expression construct
124	pPHCAGGS-H5(All Ala)	Mutant H5 expression construct
125	pPHCAGGS-GFPH5	GFP-H5 expression construct

7.2 Oligonucleotides

Name	Sequence (5'-3')	Description
H5A.for	ATCCGCTCGAGGCCATGACGGAGAGC CTGGTC	H5 cloning and sequencing
H5A.bak	GGCCGCTCGAGTTACTTCAGCTCACTT CTTCTTGGGCGATTT	H5 cloning and sequencing
H5B.for	AGGCCAAGAAGGTGAAGC	Sequencing H5 cloning junction
H5B.bak	GTAGGTGGGGTGCGATGC	Sequencing H5 cloning junction
H5C.for	GAGCCACTACAAGGTGGG	H5 sequencing
H5C.bak	TGGGCTTCTTGGCCGGTG	H5 sequencing
GFP-H5.for	ATCCGTCCGGAGGTGGCGGGATGACG GAGAGCCTGGTC	GFP-H5 cloning
BstXI/Blunt Adapter 1	GGCCGCACTGGCCAGCACA	Cloning adapter
BstXI/Blunt Adapter 2	CTGGCCAGTGCGGCC	Cloning adapter
BstXI/EcoRI Adapter 1	AATTCGGCCGCACTGGCCAGCACA	Cloning adapter
BstXI/EcoRI Adapter 2	CTGGCCACTGCGGCCG	Cloning adapter

8. References

Aagaard, L., Laible, G., Selenko, P., Schmid, M., Dorn, R., Schotta, G., Kuhfittig, S., Wolf, A., Lebersorger, A., Singh, P. B., Reuter, G., and Jenuwein, T. (1999). Functional mammalian homologues of the Drosophila PEV-modifier Su(var)3- 9 encode centromere-associated proteins which complex with the heterochromatin component M31. EMBO J. *18*, 1923-1938.

Adachi, Y., Kas, E., and Laemmli, U. K. (1989). Preferential, cooperative binding of DNA topoisomerase II to scaffold- associated regions. EMBO J. *8*, 3997-4006.

Adams, C. C. and Workman, J. L. (1993). Nucleosome Displacement In Transcription. Cell *72*, 305-308.

Adams, C. R. and Kamakaka, R. T. (1999). Chromatin assembly: biochemical identities and genetic redundancy. Curr. Opin. Genet. Dev. *9*, 185-190.

Affolter, M., Cote, J., Renaud, J., and Ruizcarrillo, A. (1987). Regulation Of Histone and Beta-a-Globin Gene-Expression During Differentiation Of Chicken Erythroid-Cells. Mol. Cell. Biol. *7*, 3663-3672.

Agarwal, S. and Rao, A. (1998). Modulation of chromatin structure regulates cytokine gene expression during T cell differentiation. Immunity. *9*, 765-775.

Agarwal, S., Viola, J. P., and Rao, A. (1999). Chromatin-based regulatory mechanisms governing cytokine gene transcription. J. Allergy Clin. Immunol. *103*, 990-999.

Ainsztein, A. M., Kandels-Lewis, S. E., Mackay, A. M., and Earnshaw, W. C. (1998). INCENP centromere and spindle targeting: identification of essential conserved motifs and involvement of heterochromatin protein HP1. J. Cell Biol. *143*, 1763-1774.

Aladjem, M. I., Spike, B. T., Rodewald, L. W., Hope, T. J., Klemm, M., Jaenisch, R., and Wahl, G. M. (1998). ES cells do not activate p53-dependent stress responses and undergo p53-independent apoptosis in response to DNA damage. Curr. Biol. *8*, 145-155.

Alexiadis, V., Varga-Weisz, P. D., Bonte, E., Becker, P. B., and Gruss, C. (1998). In vitro chromatin remodelling by chromatin accessibility complex (CHRAC) at the SV40 origin of DNA replication. EMBO J. *17*, 3428-3438.

Ali, Z. and Singh, N. (1987). Binding of linker histones to the core nucleosome. J. Biol. Chem. *262*, 12989-12993.

Allan, J., Cowling, G. J., Harborne, N., Cattini, P., Craigie, R., and Gould, H. (1981). Regulation Of the Higher-Order Structure Of Chromatin By Histones-H1 and Histones-H5. J. Cell Biol. *90*, 279-288.

Allan, J., Harborne, N., Rau, D. C., and Gould, H. (1982a). Participation of core histone "tails" in the stabilization of the chromatin solenoid. J. Cell Biol. *93*, 285-297.

Allan, J., Hartman, P. G., Crane_Robinson, C., and Aviles, F. X. (1980). The Structure of Histone H1 and its Locatation in Chromatin. Nature *288*, 675-679.

Allan, J., Mitchell, T., Harborne, N., Bohm, L., and Crane-Robinson, C. (1986). Roles of H1 Domains in Determining Higher Order Chromatin Structure and H1 Location. J. Mol. Biol. *187*, 591-601.

Allan, J., Rau, D. C., Harborne, N., and Gould, H. (1984). Higher order structure in a short repeat length chromatin. J. Cell Biol. *98*, 1320-7.

Allan, J., Smith, B. J., Dunn, B., and Bustin, M. (1982b). Antibodies against the Folding Domain of Histone H5 Cross-react with H1° But Not with H1. J. Biol. Chem. *257*, 10533-10535.

Almouzni, G., Khochbin, S., Dimitrov, S., and Wolffe, A. P. (1994). Histone acetylation influences both gene expression and development of Xenopus laevis. Dev. Biol. *165*, 654-669.

Amsterdam, A., Lin, S., and Hopkins, N. (1995). The Aequorea victoria green fluorescent protein can be used as a reporter in live zebrafish embryos. Dev. Biol. *171*, 123-129.

Annunziato, A. T., Frado, L. L., Seale, R. L., and Woodcock, C. L. (1988). Treatment with sodium butyrate inhibits the complete condensation of interphase chromatin. Chromosoma *96*, 132-138.

Annunziato, A. T. and Seale, R. L. (1983). Histone deacetylation is required for the maturation of newly replicated chromatin. J. Biol. Chem. *258*, 12675-12684.

Appels, R. and Wells, J. R. E. (1972). Synthesis and Turnover of DNA-bound Histone during Maturation of Avian Red Blood Cells. J. Mol. Biol. *70*, 425-434.

Arents, G., Burlingame, R. W., Wang, B. C., Love, W. E., and Moudrianakis, E. N. (1991). The nucleosomal core histone octamer at 3.1 A resolution: a tripartite protein assembly and a left-handed superhelix. Proc. Natl. Acad. Sci. U. S. A. *88*, 10148-10152.

Aubert, D., Garcia, M., Benchaibi, M., Poncet, D., Chebloune, Y., Verdier, G., Nigon, V., Samarut, J., and Mura, C. V. (1991). Inhibition Of Proliferation Of Primary Avian Fibroblasts Through Expression Of Histone-H5 Depends On the Degree Of Phosphorylation Of the Protein. J. Cell Biol. *113*, 497-506.

Aviles, F. J., Chapman, G. E., Kneale, G. G., Crane-Robinson, C., and Bradbury, E. M. (1978). The conformation of histone H5. Isolation and characterisation of the globular segment. Eur. J. Biochem. *88*, 363-71.

Ayer, D. E. (1999). Histone deacetylases: transcriptional repression with SINers and NuRDs. Trends. Cell Biol. *9*, 193-198.

Bain, G. and Gottlieb, D. I. (1998). Neural cells derived by in vitro differentiation of P19 and embryonic stem cells. Perspect. Dev. Neurobiol. *5*, 175-178.

Barry, A. E., Howman, E. V., Cancilla, M. R., Saffery, R., and Choo, K. H. A. (1999). Sequence analysis of an 80 kb human neocentromere. Hum. Mol. Genet. *8*, 217-227.

Baudy, P. and Bram, S. (1979). Neutron scattering on nuclei. Nucleic Acids Res. *6*, 1721-1729.

Bauer, W. R., Hayes, J. J., White, J. H., and Wolffe, A. P. (1994). Nucleosome Structural-Changes Due to Acetylation. J. Mol. Biol. *236*, 685-690.

Beato, M. and Eisfeld, K. (1997). Transcription factor access to chromatin. Nucleic Acids Res. *25*, 3559-3563.

Beddington, R. S. and Robertson, E. J. (1989). An assessment of the developmental potential of embryonic stem cells in the midgestation mouse embryo. Development *105*, 733-737.

Bednar, J., Horowitz, R. A., Grigoryev, S. A., Carruthers, L. M., Hansen, J. C., Koster, A. J., and Woodcock, C. L. (1998). Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin. Proc. Natl. Acad. Sci. U. S. A. *95*, 14173-14178.

Bednar, J., Studitsky, V. M., Grigoryev, S. A., Felsenfeld, G., and Woodcock, C. L. (1999). The nature of the nucleosomal barrier to transcription: direct observation of paused intermediates by electron cryomicroscopy. Mol. Cell *4*, 377-386.

Bell, A. C. and Felsenfeld, G. (1999). Stopped at the border: boundaries and insulators. Curr. Opin. Genet. Dev. *9*, 191-198.

Bell, A. C., West, A. G., and Felsenfeld, G. (1999). The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. Cell *98*, 387-396.

Belloc, F., Dumain, P., Boisseau, M. R., Jalloustre, C., Reiffers, J., Bernard, P., and Lacombe, F. (1994). A Flow Cytometric Method Using Hoechst 33342 and Propidium Iodide for Simultaneous Cell Cycle Analysis and Apoptosis Determination in Unfixed Cells. Cytometry *17*, 59-65.

Belmont, A. S. and Bruce, K. (1994). Visualization of G1 chromosomes: a folded, twisted, supercoiled chromonema model of interphase chromatid structure. J. Cell Biol. *127*, 287-302.

Belmont, A. S., Dietzel, S., Nye, A. C., Strukov, Y. G., and Tumbar, T. (1999). Large-scale chromatin structure and function. Curr. Opin. Cell Biol. *11*, 307-311.

Berardi, A. C., Wang, A., Levine, J. D., Lopez, P., and Scadden, D. T. (1995). Functional isolation and characterization of human hematopoietic stem cells. Science *267*, 104-108.

Berezney, R., Mortillaro, M. J., Ma, H., Wei, X., and Samarabandu, J. (1995). The nuclear matrix: a structural milieu for genomic function. Int. Rev. Cytol. *162A:1-65*, 1-65.

Bergman, M. G., Wawra, E., and Winge, M. (1988). Chicken Histone H5 Inhibits Transcription and Replication When Introduced Into Proliferating Cells By Microinjection. J. Cell Sci. *91*, 201-209.

Bernstine, E. G., Hooper, M. L., Grandchamp, S., and Ephrussi, B. (1973). Alkaline Phosphatase Activity in Mouse Teratoma. Proc. Natl. Acad. Sci. U. S. A. *70*, 3899-3903.

Bewsey, K. E., Johnson, M. E., and Huff, J. P. (1991). Rapid isolation and purification of DNA from agarose gels: the phenol- freeze-fracture method. Biotechniques *10*, 724-725.

Bickmore, W. A. and Craig, J. M. (1997). Chromosome Bands: Patterns in the Genome. (Heidelberg: Springer).

Birck, C., Poch, O., Romier, C., Ruff, M., Mengus, G., Lavigne, A. C., Davidson, I., and Moras, D. (1998). Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3. Cell *94*, 239-249.

Blank, T. A. and Becker, P. B. (1995). Electrostatic mechanism of nucleosome spacing. J. Mol. Biol. *252*, 305-313.

Blasco, M. A., Lee, H. W., Hande, M. P., Samper, E., Lansdorp, P. M., DePinho, R. A., and Greider, C. W. (1997). Telomere shortening and tumor formation by mouse cells lacking telomerase RNA. Cell *91*, 25-34.

Bloom, K. S. and Anderson, J. N. (1978). Fractionation of hen oviduct chromatin into transcriptionally active and inactive regions after selective micrococcal nuclease digestion. Cell *15*, 141-150.

Bloom, K. S. and Anderson, J. N. (1979). Conformation of ovalbumin and globin genes in chromatin during differential gene expression. J. Biol. Chem. *254*, 10532-10539.

Bosch, A. and Suau, P. (1995). Changes in core histone variant composition in differentiating neurons: the roles of differential turnover and synthesis rates. Eur. J. Cell Biol. *68*, 220-225.

Boulikas, T. (1995). Chromatin domains and prediction of MAR sequences. Int. Rev. Cytol. *162A*, 279-388.

Boulikas, T., Wiseman, J. M., and Garrard, W. T. (1980). Points of contact between histone H1 and the histone octamer. Proc. Natl. Acad. Sci. U. S. A *77*, 127-131.

Bourachot, B., Yaniv, M., and Muchardt, C. (1999). The activity of mammalian brm/SNF2α is dependent on a high-mobility- group protein I/Y-like DNA binding domain. Mol. Cell Biol. *19*, 3931-3939.

Boy de la Tour, E. and Laemmli, U. K. (1988). The metaphase scaffold is helically folded: sister chromatids have predominantly opposite helical handedness. Cell *55*, 937-944.

Bradbury, E. M. (1992). Reversible Histone Modifications and the Chromosome Cell-Cycle. Bioessays *14*, 9-16.

Bradley, A., Evans, M., Kaufman, M. H., and Robertson, E. (1984). Formation of germ-line chimaeras from embryo-derived teratocarcinoma cell lines. Nature *309*, 255-256.

Brady, G., Billia, F., Knox, J., Hoang, T., Kirsch, I. R., Voura, E. B., Hawley, R. G., Cumming, R., Buchwald, M., and Siminovitch, K. (1995). Analysis of gene expression in a complex differentiation hierarchy by global amplification of cDNA from single cells. Curr. Biol. *5*, 909-922.

Bramlage, B., Kosciessa, U., and Doenecke, D. (1997). Differential expression of the murine histone genes H3.3A and H3.3B. Differentiation *62*, 13-20.

Braunstein, M., Sobel, R. E., Allis, C. D., Turner, B. M., and Broach, J. R. (1996). Efficient transcriptional silencing in Saccharomyces cerevisiae requires a heterochromatin histone acetylation pattern. Mol. Cell Biol. *16*, 4349-4356.

Brockdorff, N. and Duthie, S. M. (1998). X Chromosome Inactivation and the Xist Gene. Cell. Mol. Life. Sci. *54*, 104-112.

Brook, F. A. and Gardner, R. L. (1997). The origin and efficient derivation of embryonic stem cells in the mouse. Proc. Natl. Acad. Sci. U. S. A. *94*, 5709-5712.

Brown, D. T., Gunjan, A., Alexander, B. T., and Sittman, D. B. (1997b). Differential effect of H1 variant overproduction on gene expression is due to differences in the central globular domain. Nucleic Acids Res. *25*, 5003-5009.

Brown, K. E., Baxter, J., Graf, D., Merkenschlager, M., and Fisher, A. G. (1999). Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. Mol. Cell *3*, 207-217.

Brown, K. E., Guest, S. S., Smale, S. T., Hahm, K., Merkenschlager, M., and Fisher, A. G. (1997a). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell *91*, 845-854.

Brown, M. T. (1995). Sequence similarities between the yeast chromosome segregation protein Mif2 and the mammalian centromere protein CENP-C. Gene *160*, 111-116.

Brown, M. T., Goetsch, L., and Hartwell, L. H. (1993). MIF2 is required for mitotic spindle integrity during anaphase spindle elongation in Saccharomyces cerevisiae. J. Cell Biol. *123*, 387-403.

Brustle, O., Jones, K. N., Learish, R. D., Karram, K., Choudhary, K., Wiestler, O. D., Duncan, I. D., and McKay, R. D. (1999). Embryonic stem cell-derived glial precursors: A source of myelinating transplants. Science *285*, 754-756.

Burdon, T., Stracey, C., Chambers, I., Nichols, J., and Smith, A. (1999). Suppression of SHP-2 and ERK signalling promotes self-renewal of mouse embryonic stem cells. Dev. Biol. *210*, 30-43.

Burns, L. J., Glauber, J. G., and Ginder, G. D. (1988). Butyrate induces selective transcriptional activation of a hypomethylated embryonic globin gene in adult erythroid cells. Blood *72*, 1536-1542.

Bustin, M. (1999). Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. Mol. Cell Biol. *19*, 5237-5246.

Bustin, M. and Reeves, R. (1996). High-mobility-group chromosomal proteins: architectural components that facilitate chromatin function. Prog. Nucleic. Acid. Res. Mol. Biol. *54:35-100*, 35-100.

Bustin, M., Trieschmann, L., and Postnikov, Y. V. (1995). The HMG-14/-17 Chromosomal Protein Family - Architectural Elements That Enhance Transcription From Chromatin Templates. Semin. Cell Biol. *6*, 247-255.

Butler, M. G., Tilburt, J., DeVries, A., Muralidhar, B., Aue, G., Hedges, L., Atkinson, J., and Schwartz, H. (1998). Comparison of chromosome telomere integrity in multiple tissues from subjects at different ages. Cancer Genet. Cytogenet. *105*, 138-144.

Butler, P. J. and Thomas, J. O. (1980). Changes in chromatin folding in solution. J. Mol. Biol. *140*, 505-529.

Callan, H. G. (1986). Lampbrush chromosomes. Mol. Biol. Biochem. Biophys. *36:1-252*, 1-252.

Calogero, S., Grassi, F., Aguzzi, A., Voigtlander, T., Ferrier, P., Ferrari, S., and Bianchi, M. E. (1999). The lack of chromosomal protein HMG1 does not disrupt cell growth but causes lethal hypoglycaemia in newborn mice. Nat. Genet. *22*, 276-280.

Candido, E. P. and Dixon, G. H. (1972). Amino-terminal sequences and sites of in vivo acetylation of trout- testis histones 3 and IIb 2. Proc. Natl. Acad. Sci. U. S. A. *69*, 2015-2019.

Caplan, A., Kimura, T., Gould, H., and Allan, J. (1987). Perturbation of chromatin structure in the region of the adult β-globin gene in chicken erythrocyte chromatin. J. Mol. Biol. *193*, 57-69.

Caplan, A. I. and Ordahl, C. P. (1978). Irreversible gene repression model for control of development. Science *201*, 120-130.

Carruthers, L. M., Bednar, J., Woodcock, C. L., and Hansen, J. C. (1998). Linker histones stabilize the intrinsic salt-dependent folding of nucleosomal arrays: mechanistic ramifications for higher-order chromatin folding. Biochemistry *37*, 14776-14787.

Cereghini, S. and Yaniv, M. (1984b). Assembly of transfected DNA into chromatin: structural changes in the origin-promoter-enhancer region upon replication. EMBO J. *3*, 1243-1253.

Chabanas, A., Lawrence, J. J., Humbert, J., and Eisen, H. (1983). Cell-Cycle Regulation Of Histone-H1° In CHO Cells - a Flow Cytofluorimetric Study After Double Staining Of the Cells. EMBO J. *2*, 833-837.

Chadee, D. N., Hendzel, M. J., Tylipski, C. P., Allis, C. D., Bazett-Jones, D. P., Wright, J. A., and Davie, J. R. (1999). Increased Ser-10 phosphorylation of histone H3 in mitogen-stimulated and oncogene-transformed mouse fibroblasts. J. Biol. Chem. *274*, 24914-24920.

Chalfie, M., Tu, Y., Euskirchen, G., Ward, W. W., and Prasher, D. C. (1994). Green Fluorescent Protein As a Marker For Gene-Expression. Science *263*, 802-805.

Chambers, I., Cozens, A., Broadbent, J., Robertson, M., Lee, M., Li, M., and Smith, A. (1997). Structure of the mouse leukaemia inhibitory factor receptor gene: regulated expression of mRNA encoding a soluble receptor isoform from an alternative 5' untranslated region. Biochem. J. *328*, 879-888.

Chang, L., Loranger, S. S., Mizzen, C., Ernst, S. G., Allis, C. D., and Annunziato, A. T. (1997). Histones in transit: cytosolic histone complexes and diacetylation of H4 during nucleosome assembly in human cells. Biochemistry *36*, 469-480.

Chen, D., Ma, H., Hong, H., Koh, S. S., Huang, S. M., Schurter, B. T., Aswad, D. W., and Stallcup, M. R. (1999). Regulation of transcription by a protein methyltransferase. Science *284*, 2174-2177.

Cherry, S. R. and Baltimore, D. (1999). Chromatin remodeling directly activates V(D)J recombination. Proc. Natl. Acad. Sci. U. S. A *96*, 10788-10793.

Chittenden, T., Frey, A., and Levine, A. J. (1991). Regulated Replication of an Episomal Simian Virus 40 Origin Plasmid in COS7 Cells. J. Virol. *65*, 5944-5951.

Cho, S. K., Webber, T. D., Carlyle, J. R., Nakano, T., Lewis, S. M., and Zuniga-Pflucker, J. C. (1999). Functional characterization of B lymphocytes generated in vitro from embryonic stem cells. Proc. Natl. Acad. Sci. U. S. A. *96*, 9797-9802.

Chomczynski, P. and Sacchi, N. (1987). Single-step method of RNA isolation by acid guanidinium thiocyanate- phenol-chloroform extraction. Anal. Biochem. *162*, 156-159.

Choo, K. H., Vissel, B., and Earle, E. (1989). Evolution of α-Satellite DNA On Human Acrocentric Chromosomes. Genomics *5*, 332-344.

Choo, K. H. A. (1997). The Centromere. (Oxford: Oxford University Press).

Church, G. M. and Gilbert, W. (1984). Genomic Sequencing. Proc. Natl. Acad. Sci. U. S. A. *81*, 1991-1995.

Clark, D. J. and Kimura, T. (1990). Electrostatic Mechanism of Chromatin Folding. J. Mol. Biol. *211*, 883-896.

Clark, K. L., Halay, E. D., Lai, E. S., and Burley, S. K. (1993). Co-Crystal Structure of the Hnf-3/Fork Head DNA-Recognition Motif Resembles Histone-H5. Nature *364*, 412-420.

Clarke, H. J., McLay, D. W., and Mohamed, O. A. (1998). Linker histone transitions during mammalian oogenesis and embryogenesis. Dev. Genet. *22*, 17-30.

Cleard, F., Delattre, M., and Spierer, P. (1997). SU(VAR)3-7, a Drosophila heterochromatin-associated protein and companion of HP1 in the genomic silencing of position-effect variegation. EMBO J. *16*, 5280-5288.

Clemson, C. M., McNeil, J. A., Willard, H. F., and Lawrence, J. B. (1996). XIST RNA paints the inactive X chromosome at interphase: evidence for a novel RNA involved in nuclear/chromosome structure. J. Cell Biol. *132*, 259-275.

Clore, G. M., Gronenborn, A. M., Nilges, M., Sukumaran, D. K., and Zarbock, J. (1987). The polypeptide fold of the globular domain of histone H5 in solution. A study using nuclear magnetic resonance, distance geometry and restrained molecular dynamics. EMBO J. *6*, 1833-1842.

Corbet, S. W., Clarke, A. R., Gledhill, S., and Wyllie, A. H. (1999). P53-dependent and -independent links between DNA-damage, apoptosis and mutation frequency in ES cells. Oncogene *18*, 1537-1544.

Cormack, B. P., Valdivia, R. H., and Falkow, S. (1996). FACS-Optimized Mutants of the Green Fluorescent Protein (GFP). Gene *173*, 33-38.

Cosma, M. P., Tanaka, T., and Nasmyth, K. (1999). Ordered recruitment of transcription and chromatin remodeling factors to a cell cycle- and developmentally regulated promoter. Cell *97*, 299-311.

Costanzi, C. and Pehrson, J. R. (1998). Histone macroH2A1 is concentrated in the inactive X chromosome of female mammals. Nature *393*, 599-601.

Craig, J. M., Earnshaw, W. C., and Vagnarelli, P. (1999). Mammalian centromeres: DNA sequence, protein composition, and role in cell cycle progression. Exp. Cell Res. *246*, 249-262.

Crane-Robinson, C. (1997). Where is the globular domain of linker histone located on the nucleosome? Trends Biochem. Sci. *22* , 75-77.

Crane-Robinson, C. (1999). How do linker histones mediate differential gene expression? Bioessays *21*, 367-371.

Cross, S. H., Meehan, R. R., Nan, X., and Bird, A. (1997). A component of the transcriptional repressor MeCP1 shares a motif with DNA methyltransferase and HRX proteins. Nat. Genet. *16*, 256-259.

Cryderman, D. E., Morris, E. J., Biessmann, H., Elgin, S. C., and Wallrath, L. L. (1999). Silencing at Drosophila telomeres: nuclear organization and chromatin structure play critical roles. EMBO J. *18*, 3724-3735.

Csankovszki, G., Panning, B., Bates, B., Pehrson, J. R., and Jaenisch, R. (1999). Conditional deletion of Xist disrupts histone macroH2A localization but not maintenance of X inactivation. Nat. Genet. *22*, 323-324.

Csink, A. K. and Henikoff, S. (1996). Genetic modification of heterochromatic association and nuclear organization in Drosophila. Nature *381*, 529-531.

Csink, A. K. and Henikoff, S. (1998). Something from nothing: The evolution and utility of satellite repeats. Trends Genet. *14*, 200-204.

Cubitt, A. B., Heim, R., Adams, S. R., Boyd, A. E., Gross, L. A., and Tsien, R. Y. (1995). Understanding, Improving and Using Green Fluorescent Proteins. Trends Biochem. Sci. *20*, 448-455.

Cutts, S. M., Fowler, K. J., Kile, B. T., Hii, L. L., O'Dowd, R. A., Hudson, D. F., Saffery, R., Kalitsis, P., Earle, E., and Choo, K. H. (1999). Defective chromosome segregation, microtubule bundling and nuclear bridging in inner centromere protein gene (Incenp)-disrupted mice. Hum. Mol. Genet. *8*, 1145-1155.

Cuvier, O., Hart, C. M., and Laemmli, U. K. (1998). Identification of a class of chromatin boundary elements. Mol. Cell Biol. *18*, 7478-7486.

D'Anna, J. A., Strniste, G. F., and Gurley, L. R. (1979). Circular dichroic and sedimentation studies of phosphorylated H1 from Chinese hamster cells. Biochemistry *18*, 942-951.

Dani, C., Smith, A. G., Dessolin, S., Leroy, P., Staccini, L., Villageois, P., Darimont, C., and Ailhaud, G. (1997). Differentiation of embryonic stem cells into adipocytes in vitro. J. Cell Sci. *110*, 1279-1285.

Davey, C., Pennings, S., and Allan, J. (1997). CpG methylation remodels chromatin structure in vitro. J. Mol. Biol. *267*, 276-288.

Davey, C., Pennings, S., Meersseman, G., Wess, T. J., and Allan, J. (1995). Periodicity of Strong Nucleosome Positioning Sites Around the Chicken Adult Beta-Globin Gene May Encode Regularly Spaced Chromatin. Proc. Natl. Acad. Sci. U. S. A. *92*, 11210-11214.

Davie, J. R. and Chadee, D. N. (1998). Regulation and regulatory parameters of histone modifications. J. Cell Biochem. Suppl. *30-31:203-13*, 203-213.

Deacon, T., Dinsmore, J., Costantini, L. C., Ratliff, J., and Isacson, O. (1998). Blastula-stage stem cells can differentiate into dopaminergic and serotonergic neurons after transplantation. Exp. Neurol. *149*, 28-41.

DeLange, R. J., Fambrough, D. M., Smith, E. L., and Bonner, J. (1969). Calf and pea histone IV. 3. Complete amino acid sequence of pea seedling histone IV; comparison with the homologous calf thymus histone. J. Biol. Chem. *244*, 5669-5679.

Dernburg, A. F., Broman, K. W., Fung, J. C., Marshall, W. F., Philips, J., Agard, D. A., and Sedat, J. W. (1996a). Perturbation of nuclear architecture by long-distance chromosome interactions. Cell *85*, 745-759.

Dernburg, A. F., Sedat, J. W., and Hawley, R. S. (1996b). Direct evidence of a role for heterochromatin in meiotic chromosome segregation. Cell *86*, 135-146.

Dietzel, S., Niemann, H., Bruckner, B., Maurange, C., and Paro, R. (1999). The nuclear distribution of Polycomb during Drosophila melanogaster development shown with a GFP fusion protein. Chromosoma *108*, 83-94.

Dimova, D., Nackerdien, Z., Furgeson, S., Eguchi, S., and Osley, M. A. (1999). A role for transcriptional repressors in targeting the yeast Swi/Snf complex. Mol. Cell *4*, 75-83.

Ding, H. F., Bustin, M., and Hansen, U. (1997). Alleviation of histone H1-mediated transcriptional repression and chromatin compaction by the acidic activation region in chromosomal protein HMG-14. Mol. Cell Biol. *17*, 5843-5855.

Dinsmore, J., Ratliff, J., Deacon, T., Pakzaban, P., Jacoby, D., Galpern, W., and Isacson, O. (1996). Embryonic Stem-Cells Differentiated In-Vitro As a Novel Source Of Cells For Transplantation. Cell Transplant. *5*, 131-143.

Divane, A., Carter, N. P., Spathas, D. H., and Ferguson-Smith, M. A. (1994). Rapid prenatal diagnosis of aneuploidy from uncultured amniotic fluid cells using five-colour fluorescence in situ hybridization. Prenat. Diagn. *14*, 1061-1069.

Doetschman, T. C., Eistetter, H., Katz, M., Schmidt, W., and Kemler, R. (1985). The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. J. Embryol. Exp. Morphol. *87:27-45*, 27-45.

Dolby, T. N., Ajiro, K., Borun, T. W., Gilmour, R. S., Zweidler, A., Cohen, L., Miller, P., and Nieolini, C. (1979). Physical properties of DNA and chromatin isolated from G1- and S-phase HeLa S-3 cells. Effects of histone H1 phosphorylation and stage- specific nonhistone chromosomal proteins on the molar ellipticity of native and reconstituted nucleoproteins during thermal denaturation. Biochemistry *18*, 1333-44.

DuPraw, E. J. (1968). Organisation of Genetic Material in Eukaryotic Chromosomes. In Cell and Molecular Biology, (New York: Academic Press), pp. 514-589.

Eichler, E. E. (1999). Repetitive conundrums of centromere structure and function. Hum. Mol. Genet. *8*, 151-155.

Eissenberg, J. C., Ge, Y. W., and Hartnett, T. (1994). Increased phosphorylation of HP1, a heterochromatin-associated protein of Drosophila, is correlated with heterochromatin assembly. J. Biol. Chem. *269* , 21315-21321.

Evans, E. P. (1981). Karyotype of the house mouse. In Biology of the house mouse, R.J.Berry, ed., pp. 127-139.

Evans, M. J. and Kaufman, M. H. (1981). Establishment in culture of pluripotential cells from mouse embryos. Nature *292*, 154-156.

Faria, T. N., Mendelsohn, C., Chambon, P., and Gudas, L. J. (1999). The Targeted Disruption of Both Alleles of RARβ(2) in F9 Cells Results in the Loss of Retinoic Acid-associated Growth Arrest. J. Biol. Chem. *274*, 26783-26788.

Feinberg, A. P. and Vogelstein, B. (1983). A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochem. *132*, 6-13.

Felsenfeld, G. and McGhee, J. D. (1986). Structure of the 30 nm chromatin fiber. Cell *44*, 375-7.

Figueroa, J., Saffrich, R., Ansorge, W., and Valdivia, M. (1998). Microinjection of antibodies to centromere protein CENP-A arrests cells in interphase but does not prevent mitosis. Chromosoma *107*, 397-405.

Filipski, J., Leblanc, J., Youdale; T., Sikorska, M., and Walker, P. R. (1990b). Periodicity of DNA folding in higher order chromatin structures. EMBO J. *9*, 1319-1327.

Finch, J. T. and Klug, A. (1976). Solenoid model for superstructure in chromatin. Proc. Natl. Acad. Sci. U. S. A. *73*, 1897-1901.

Fisher, E. A. and Felsenfeld, G. (1986). Comparison of the folding of β-globin and ovalbumin gene containing chromatin isolated from chicken oviduct and erythrocytes. Biochemistry *25*, 8010-8016.

Fittler, F. and Zachau, H. G. (1979). Subunit structure of α-satellite DNA containing chromatin from African green monkey cells. Nucleic Acids Res. *7*, 1-13.

Fletcher, T. M. and Hansen, J. C. (1995). Core histone tail domains mediate oligonucleosome folding and nucleosomal DNA organization through distinct molecular mechanisms. J. Biol. Chem. *270*, 25359-25362.

Fletcher, T. M. and Hansen, J. C. (1996). The nucleosomal array: structure/function relationships. Crit. Rev. Eukaryot. Gene Express. *6*, 149-188.

Forster, A., Buluwela, L., and Rabbitts, T. H. (1990). Turbo-screening of bacterial colonies using microwave denaturation on paper filters. Trends. Genet. *6*, 141.

Fraichard, A., Chassande, O., Bilbaut, G., Dehay, C., Savatier, P., and Samarut, J. (1995). In-Vitro Differentiation of Embryonic Stem-Cells Into Glial-Cells and Functional-Neurons. J. Cell Sci. *108*, 3181-3188.

Fujimura, F. K., Deininger, P. L., Friedmann, T., and Linney, E. (1981). Mutation near the Polyoma DNA Replication Origin Permits Productive Infection of F9 Embryonal Carcinoma Cell. Cell *23*, 809-814.

Fukagawa, T. and Brown, W. R. (1997). Efficient conditional mutation of the vertebrate CENP-C gene. Hum. Mol. Genet. *6*, 2301-2308.

Fukagawa, T., Pendon, C., Morris, J., and Brown, W. (1999). CENP-C is necessary but not sufficient to induce formation of a functional centromere. EMBO J. *18*, 4196-4209.

Gaff, C., du, S. D., Kalitsis, P., Iannello, R., Nagy, A., and Choo, K. H. (1994). A novel nuclear protein binds centromeric α -satellite DNA. Hum. Mol. Genet. *3*, 711-716.

Gaiano, N., Allende, M., Amsterdam, A., Kawakami, K., and Hopkins, N. (1996). Highly efficient germ-line transmission of proviral insertions in zebrafish. Proc. Natl. Acad. Sci. U. S. A. *93*, 7777-7782.

Garcia-Ramirez, M., Dong, F., and Ausio, J. (1992). Role of the histone "tails" in the folding of oligonucleosomes depleted of histone H1. J. Biol. Chem. *267*, 19587-19595.

Garcia-Ramirez, M., Rocchini, C., and Ausio, J. (1995). Modulation of chromatin folding by histone acetylation. J. Biol. Chem. *270*, 17923-17928.

Gassmann, M., Donoho, G., and Berg, P. (1995). Maintenance Of an Extrachromosomal Plasmid Vector In Mouse Embryonic Stem-Cells. Proc. Natl. Acad. Sci. U. S. A. *92*, 1292-1296.

Gaszner, M., Vazquez, J., and Schedl, P. (1999). The Zw5 protein, a component of the scs chromatin domain boundary, is able to block enhancer-promoter interaction. Genes Dev. *13*, 2098-2107.

Georgel, P. T., Tsukiyama, T., and Wu, C. (1997). Role of histone tails in nucleosome remodeling by Drosophila NURF. EMBO J. *16*, 4717-4726.

Gerchman, S. E., Graziano, V., and Ramakrishnan, V. (1994). Expression of Chicken Linker Histones in Escherichia-Coli - Sources of Problems and Methods For Overcoming Some of the Difficulties. Protein Expr. Purif. *5*, 242-251.

Germond, J. E., Bellard, M., Oudet, P., and Chambon, P. (1976). Stability of nucleosomes in native and reconstituted chromatins. Nucleic Acids Res. *3*, 3173-3192.

Giancotti, V., Bandiera, A., Ciani, L., Santoro, D., Cranerobinson, C., Goodwin, G. H., Boiocchi, M., Dolcetti, R., and Casetta, B. (1993). High-Mobility-Group (HMG) Proteins and Histone H1 Subtypes Expression In Normal and Tumor-Tissues Of Mouse. Eur. J. Biochem. *213*, 825-832.

Girardot, V., Rabilloud, T., Yoshida, M., Beppu, T., Lawrence, J. J., and Khochbin, S. (1994). Relationship Between Core Histone Acetylation and Histone H1° Gene Activity. Eur. J. Biochem. *224*, 885-892.

Gjerset, R., Gorka, C., Hasthorpe, S., Lawrence, J. J., and Eisen, H. (1982). Developmental and Hormonal-Regulation Of Protein H1° In Rodents. Proc. Natl. Acad. Sci. U. S. A. *79*, 2333-2337.

Glauber, J. G., Wandersee, N. J., Little, J. A., and Ginder, G. D. (1991). 5'-flanking sequences mediate butyrate stimulation of embryonic globin gene expression in adult erythroid cells. Mol. Cell Biol. *11*, 4690-4697.

Goldberg, M., Harel, A., Brandeis, M., Rechsteiner, T., Richmond, T. J., Weiss, A. M., and Gruenbaum, Y. (1999). The tail domain of lamin Dm0 binds histones H2A and H2B. Proc. Natl. Acad. Sci. U. S. A *96*, 2852-2857.

Gorka, C., Fakan, S., and Lawrence, J. J. (1993). Light and Electron-Microscope Immunocytochemical Analyses Of Histone H1° Distribution In the Nucleus Of Friend-Erythroleukemia Cells. Exp. Cell Res. *205*, 152-158.

Gossen, M. and Bujard, H. (1992). Tight control of gene expression in mammalian cells by tetracycline- responsive promoters. Proc. Natl. Acad. Sci. U. S. A. *89*, 5547-5551.

Gossen, M., Freundlieb, S., Bender, G., Muller, G., Hillen, W., and Bujard, H. (1995). Transcriptional Activation By Tetracyclines In Mammalian-Cells. Science *268*, 1766-1769.

Goto, H., Tomono, Y., Ajiro, K., Kosako, H., Fujita, M., Sakurai, M., Okawa, K., Iwamatsu, A., Okigaki, T., Takahashi, T., and Inagaki, M. (1999). Identification of a novel phosphorylation site on histone H3 coupled with mitotic chromosome condensation. J. Biol. Chem. *274*, 25543-25549.

Gottesfeld, J. M., Murphy, R. F., and Bonner, J. (1975). Structure of transcriptionally active chromatin. Proc. Natl. Acad. Sci. U. S. A. *72*, 4404-4408.

Goytisolo, F. A., Gerchman, S. E., Yu, X., Rees, C., Graziano, V., Ramakrishnan, V., and Thomas, J. O. (1996). Identification of two DNA-binding sites on the globular domain of histone H5. EMBO J. *15*, 3421-3429.

Graziano, V., Gerchman, S. E., and Ramakrishnan, V. (1988). Reconstitution of chromatin higher-order structure from histone H5 and depleted chromatin. J. Mol. Biol. *203*, 997-1007.

Graziano, V., Gerchman, S. E., Schneider, D. K., and Ramakrishnan, V. (1994). Histone H1 Is Located in the Interior of the Chromatin 30nm Filament. Nature *368*, 351-354.

Greider, C. W. (1996). Telomere length regulation. Annu. Rev. Biochem. *65*, 337-365.

Griffith, J. D., Comeau, L., Rosenfield, S., Stansel, R. M., Bianchi, A., Moss, H., and de Lange, T. (1999). Mammalian telomeres end in a large duplex loop. Cell *97*, 503-514.

Grigoryev, S. A., Bednar, J., and Woodcock, C. L. (1999). MENT, a heterochromatin protein that mediates higher order chromatin folding, is a new serpin family member. J. Biol. Chem. *274*, 5626-5636.

Grigoryev, S. A. and Woodcock, C. L. (1993). Stage-specific expression and localization of MENT, a nuclear protein associated with chromatin condensation in terminally differentiating avian erythroid cells. Exp. Cell Res. *206*, 335-343.

Groft, C. M., Uljon, S. N., Wang, R., and Werner, M. H. (1998). Structural homology between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly. Proc. Natl. Acad. Sci. U. S. A. *95*, 9117-9122.

Grosveld, F. (1999). Activation by locus control regions? Curr. Opin. Genet. Dev. *9*, 152-157.

Grosveld, F., Blom van Assendelft, G., Greaves, D. R., and Kollias, G. (1987). Position-independant, high-level expression of the human α-globin gene in transgenic mice. Cell *51*, 975-985.

Grozinger, C. M., Hassig, C. A., and Schreiber, S. L. (1999). Three proteins define a class of human histone deacetylases related to yeast Hda1p. Proc. Natl. Acad. Sci. U. S. A. *96*, 4868-4873.

Gunjan, A. and Brown, D. T. (1999). Overproduction of histone H1 variants *in vivo* increases basal and induced activity of the mouse mammary tumor virus promoter. Nucleic Acids Res. *27*, 3355-3363.

Guyon, J. R., Narlikar, G. J., Sif, S., and Kingston, R. E. (1999). Stable remodeling of tailless nucleosomes by the human SWI-SNF complex. Mol. Cell Biol. *19*, 2088-2097.

Haaf, T., Warburton, P. E., and Willard, H. F. (1992). Integration of human α-satellite DNA into simian chromosomes: centromere protein binding and disruption of normal chromosome segregation. Cell *70*, 681-696.

Hahm, K., Cobb, B. S., McCarty, A. S., Brown, K. E., Klug, C. A., Lee, R., Akashi, K., Weissman, I. L., Fisher, A. G., and Smale, S. T. (1998). Helios, a T cell-restricted Ikaros family member that quantitatively associates with Ikaros at centromeric heterochromatin. Genes Dev. *12*, 782-796.

Halmer, L. and Gruss, C. (1996). Effects Of Cell-Cycle Dependent Histone H1 Phosphorylation On Chromatin Structure and Chromatin Replication. Nucleic Acids Res. *24*, 1420-1427.

Han, M., Chang, M., Kim, U. J., and Grunstein, M. (1987). Histone H2B repression causes cell-cycle-specific arrest in yeast: effects on chromosomal segregation, replication, and transcription. Cell *48*, 589-597.

Hanahan, D. (1983). Studies on transformation of Escherichia coli with plasmids. J. Mol. Biol. *166*, 557-580.

Hansen, J. C., Tse, C., and Wolffe, A. P. (1998). Structure and function of the core histone N-termini: more than meets the eye. Biochemistry *37*, 17637-17641.

Harley, C. B., Futcher, A. B., and Greider, C. W. (1990). Telomeres shorten during ageing of human fibroblasts. Nature *345*, 458-460.

Harrington, J. J., VanBokkelen, G., Mays, R. W., Gustashaw, K., and Willard, H. F. (1997). Formation of de novo centromeres and construction of first-generation human artificial microchromosomes. Nat. Genet. *15*, 345-355.

Hart, C. M. and Laemmli, U. K. (1998). Facilitation of chromatin dynamics by SARs. Curr. Opin. Genet. Dev. *8*, 519-525.

Hart, C. M., Zhao, K., and Laemmli, U. K. (1997). The scs' boundary element: characterization of boundary element- associated factors. Mol. Cell Biol. *17*, 999-1009.

Hartman, P. G., Chapman, G. E., Moss, T., and Bradbury, E. M. (1977). Studies on the role and mode of operation of the very-lysine-rich histone H1 in eukaryote chromatin. The three structural regions of the histone H1 molecule. Eur. J. Biochem. *77*, 45-51.

Hayes, J. J. (1996). Site-directed cleavage of DNA by a linker histone--Fe(II) EDTA conjugate: localization of a globular domain binding site within a nucleosome. Biochemistry *35*, 11931-11937.

Hayes, J. J. and Wolffe, A. P. (1993). Preferential and asymmetric interaction of linker histones with 5S DNA in the nucleosome. Proc. Natl. Acad. Sci. U. S. A. *90*, 6415-6419.

He, D. C., Zeng, C. Q., Woods, K., Zhong, L., Turner, D., Busch, R. K., Brinkley, B. R., and Busch, H. (1998). CENP-G: a new centromeric protein that is associated with the α-1 satellite DNA subfamily. Chromosoma *107*, 189-197.

Hebbes, T. R., Thorne, A. W., Clayton, A. L., and Crane-Robinson, C. (1992). Histone acetylation and globin gene switching. Nucleic Acids Res. *20*, 1017-1022.

Hebbes, T. R., Thorne, A. W., and Crane-Robinson, C. (1988). A direct link between core histone acetylation and transcriptionally active chromatin. EMBO J. *7*, 1395-1402.

Hendrich, B. and Bird, A. (1998). Identification and characterization of a family of mammalian methyl-CpG binding proteins. Mol. Cell Biol. *18*, 6538-6547.

Hennig, W. (1999). Heterochromatin. Chromosoma *108*, 1-9.

Herrera, E., Samper, E., Caballero, J., Flores, J. M., Lee, H. W., and Blasco, M. A. (1999b). Disease states associated with telomerase deficiency appear earlier in mice with short telomeres. EMBO J. *18*, 2950-2960.

Herrera, J. E., Sakaguchi, K., Bergel, M., Trieschmann, L., Nakatani, Y., and Bustin, M. (1999a). Specific acetylation of chromosomal protein HMG-17 by PCAF alters its interaction with nucleosomes. Mol. Cell Biol. *19*, 3466-3473.

Heyworth, C. M., Alauldin, M., Cross, M. A., Fairbairn, L. J., Dexter, T. M., and Whetton, A. D. (1995). Erythroid development of the FDCP-Mix A4 multipotent cell line is governed by the relative concentrations of erythropoietin and interleukin 3. Br. J. Haematol. *91*, 15-22.

Heyworth, C. M., Dexter, T. M., Kan, O., and Whetton, A. D. (1990). The role of hemopoietic growth factors in self-renewal and differentiation of IL-3-dependent multipotential stem cells. Growth Factors *2* , 197-211.

Higashijima, S., Okamoto, H., Ueno, N., Hotta, Y., and Eguchi, G. (1997). High-frequency generation of transgenic zebrafish which reliably express GFP in whole muscles or the whole body by using promoters of zebrafish origin. Dev. Biol. *192*, 289-299.

Higgs, D. R. (1998). Do LCRs Open Chromatin Domains? Cell *95*, 299-302.

Hill, C. S., Rimmer, J. M., Green, B. N., Finch, J. T., and Thomas, J. O. (1991). Histone-DNA interactions and their modulation by phosphorylation of - Ser-Pro-X-Lys/Arg- motifs. EMBO J. *10*, 1939-1948.

Hock, R., Wilde, F., Scheer, U., and Bustin, M. (1998). Dynamic relocation of chromosomal protein HMG-17 in the nucleus is dependent on transcriptional activity. EMBO J. *17*, 6992-7001.

Hoffmann, A., Chiang, C. M., Oelgeschlager, T., Xie, X. L., Burley, S. K., Nakatani, Y., and Roeder, R. G. (1996). A Histone Octamer-Like Structure Within TFIID. Nature *380*, 356-359.

Hogan, B. L. M., Taylor, A., and Adamson, E. (1981). Cell Interactions Modulate Embryonal Carcinoma Cell Differentiation into Parietal or Visceral Endoderm. Nature *291*, 235-237.

Hohmann, P. (1983). Phosphorylation of H1 histones. Mol. Cell Biochem. *57*, 81-92.

Hole, N., Graham, G. J., Menzel, U., and Ansell, J. D. (1996). A limited temporal window for the derivation of multilineage repopulating hematopoietic progenitors during embryonal stem cell differentiation in vitro. Blood *88*, 1266-1276.

Horowitz, R. A., Agard, D. A., Sedat, J. W., and Woodcock, C. L. (1994). The three-dimensional architecture of chromatin in situ: electron tomography reveals fibers composed of a continuously variable zig-zag nucleosomal ribbon. J. Cell Biol. *125*, 1-10.

Horowitz, R. A., Koster, A. J., Walz, J., and Woodcock, C. L. (1997). Automated electron microscope tomography of frozen-hydrated chromatin: the irregular three-dimensional zigzag architecture persists in compact, isolated fibers. J. Struct. Biol. *120*, 353-362.

Horsley, D., Hutchings, A., Butcher, G. W., and Singh, P. B. (1996). M32, a murine homologue of Drosophila heterochromatin protein 1 (HP1), localises to euchromatin within interphase nuclei and is largely excluded from constitutive heterochromatin. Cytogenet. Cell Genet. *73*, 308-311.

Horvath, P. and Horz, W. (1981). The compaction of mouse heterochromatin as studied by nuclease digestion. FEBS Lett. *134*, 25-28.

Horz, W. and Altenburger, W. (1981). Nucleotide Sequence of Mouse Satellite DNA. Nucleic Acids Res. *9*, 683-696.

Horz, W., Fittler, F., and Zachau, H. G. (1983). Sequence Specific Cleavage of African-Green Monkey Alpha-Satellite DNA By Micrococcal Nuclease. Nucleic Acids Res. *11*, 4275-4285.

Horz, W. and Zachau, H. G. (1977). Characterization of distinct segments in mouse satellite DNA by restriction nucleases. Eur. J. Biochem. *73*, 383-92.

Hozak, P., Cook, P. R., Schofer, C., Mosgoller, W., and Wachtler, F. (1994). Site of Transcription of Ribosomal-RNA and Intranucleolar Structure in Hela-Cells. J. Cell Sci. *107*, 639-648.

Hozier, J., Renz, M., and Nehls, P. (1977). The chromosome fiber: evidence for an ordered superstructure of nucleosomes. Chromosoma *62*, 301-317.

Hu, M., Krause, D., Greaves, M., Sharkis, S., Dexter, M., Heyworth, C., and Enver, T. (1997). Multilineage gene expression precedes commitment in the hemopoietic system. Genes Dev. *11*, 774-785.

Huang, H., Smothers, J. F., Wiley, E. A., and Allis, C. D. (1999). A nonessential HP1-like protein affects starvation-induced assembly of condensed chromatin and gene expression in macronuclei of Tetrahymena thermophila. Mol. Cell Biol. *19*, 3624-3634.

Huang, H., Wiley, E. A., Lending, C. R., and Allis, C. D. (1998). An HP1-like protein is missing from transcriptionally silent micronuclei of Tetrahymena. Proc. Natl. Acad. Sci. U. S. A. *95*, 13624-13629.

Hudson, D. F., Fowler, K. J., Earle, E., Saffery, R., Kalitsis, P., Trowell, H., Hill, J., Wreford, N. G., de Kretser, D. M., Cancilla, M. R., Howman, E., Hii, L., Cutts, S. M., Irvine, D. V., and Choo, K. H. (1998). Centromere protein B null mice are mitotically and meiotically normal but have lower body and testis weights. J. Cell Biol. *141*, 309-319.

Huebner, K., Linnenbach, A., Weidner, S., Glenn, G., and Croce, C. M. (1981). Deoxyribonuclease-I Sensitivity Of Plasmid Genomes In Teratocarcinoma-Derived Stem and Differentiated Cells. Proc. Natl. Acad. Sci. U. S. A. *78*, 5071-5075.

Igo-Kemenes, T., Greil, W., and Zachau, H. G. (1977). Prepartation of soluble chromatin and specific chromatin fractions with restriction nucleases. Nucleic Acids Res. *4*, 3387-3400.

Ikeda, K., Steger, D. J., Eberharter, A., and Workman, J. L. (1999). Activation domain-specific and general transcription stimulation by native histone acetyltransferase complexes. Mol. Cell Biol. *19*, 855-863.

Ito, T., Bulger, M., Kobayashi, R., and Kadonaga, J. T. (1996). Drosophila NAP-1 is a core histone chaperone that functions in ATP- facilitated assembly of regularly spaced nucleosomal arrays. Mol. Cell Biol. *16*, 3112-3124.

Ivarie, R. D., Schacter, B. S., and O'Farrell, P. H. (1983). The level of expression of the rat growth hormone gene in liver tumor cells is at least eight orders of magnitude less than that in anterior pituitary cells. Mol. Cell. Biol. *3*, 1460-1467.

Izaurralde, E., Kas, E., and Laemmli, U. K. (1989). Highly preferential nucleation of histone H1 assembly on scaffold- associated regions. J. Mol. Biol. *210*, 573-585.

Jackson, D. A. and Cook, P. R. (1985). Transcription occurs at a nucleoskeleton. EMBO J. *4*, 919-925.

Jackson, M. S., Rocchi, M., Thompson, G., Hearn, T., Crosier, M., Guy, J., Kirk, D., Mulligan, L., Ricco, A., Piccininni, S., Marzella, R., Viggiano, L., and Archidiacono, N. (1999). Sequences flanking the centromere of human chromosome 10 are a complex patchwork of arm-specific sequences, stable duplications and unstable sequences with homologies to telomeric and other centromeric locations. Hum. Mol. Genet. *8*, 205-215.

Jackson, V., Marshall, S., and Chalkley, R. (1981). The sites of deposition of newly synthesized histone. Nucleic Acids Res. *9*, 4563-4581.

James, T. C. and Elgin, S. C. R. (1986). Identification of a Nonhistone Chromosomal Protein Associated With Heterochromatin in Drosophila-Melanogaster and Its Gene. Mol. Cell. Biol. *6*, 3862-3872.

Jasinskas, A. and Hamkalo, B. A. (1999). Purification and initial characterization of primate satellite chromatin. Chromosome Res. *7*, 341-354.

Jayaraman, L., Moorthy, N. C., Murthy, K. G., Manley, J. L., Bustin, M., and Prives, C. (1998). High mobility group protein-1 (HMG-1) is a unique activator of p53. Genes Dev. *12*, 462-72.

Jimenez, G., Griffiths, S. D., Ford, A. M., Greaves, M. F., and Enver, T. (1992). Activation Of the Beta-Globin Locus-Control Region Precedes Commitment to the Erythroid Lineage. Proc. Natl. Acad. Sci. U. S. A. *89*, 10618-10622.

Johnson, K. R., Disney, J. E., Wyatt, C. R., and Reeves, R. (1990). Expression Of Messenger-RNAs Encoding Mammalian Chromosomal-Proteins HMG-I and HMG-Y During Cellular Proliferation. Exp. Cell Res. *187*, 69-76.

Johnson, K. R., Lehn, D. A., Elton, T. S., Barr, P. J., and Reeves, R. (1988). Complete Murine cDNA Sequence, Genomic Structure, and Tissue Expression Of the High Mobility Group Protein HMG-I(Y). J. Biol. Chem. *263*, 18338-18342.

Johnson, K. R., Lehn, D. A., and Reeves, R. (1989). Alternative Processing Of Messenger-RNAs Encoding Mammalian Chromosomal High-Mobility-Group Proteins HMG-I and HMG-Y. Mol. Cell. Biol. *9*, 2114-2123.

Jones, K. W. (1970). Chromosomal and nuclear location of mouse satellite DNA in individual cells. Nature *225*, 912-5.

Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolffe, A. P. (1998). Methylated DNA and MeCP2 recruit histone deacetylase to repress transcription. Nat. Genet. *19*, 187-191.

Joseph, A., Mitchell, A. R., and Miller, O. J. (1989). The Organization of the Mouse Satellite DNA At Centromeres. Exp. Cell Res. *183*, 494-500.

Juan, G., Traganos, F., and Darzynkiewicz, Z. (1999). Histone H3 phosphorylation in human monocytes and during HL-60 cell differentiation. Exp. Cell Res. *246*, 212-220.

Kalitsis, P., Fowler, K. J., Earle, E., Hill, J., and Choo, K. H. A. (1998a). Targeted disruption of mouse centromere protein C gene leads to mitotic disarray and early embryo death. Proc. Natl. Acad. Sci. U. S. A. *95*, 1136-1141.

Kalitsis, P., MacDonald, A., Newson, A. J., Hudson, D. F., and Choo, K. H. A. (1998b). Gene structure and sequence analysis of mouse centromere proteins A and C. Genomics *47*, 108-114.

Kamakaka, R. T. and Thomas, J. O. (1990). Chromatin Structure of Transcriptionally Competent and Repressed Genes. EMBO J. *9*, 3997-4006.

Kanda, T., Sullivan, K. F., and Wahl, G. M. (1998). Histone-GFP fusion protein enables sensitive analysis of chromosome dynamics in living mammalian cells. Curr. Biol. *8*, 377-85.

Kapoor, M., Montes de Oca, L. R., Liu, G., Lozano, G., Cummings, C., Mancini, M., Ouspenski, I., Brinkley, B. R., and May, G. S. (1998). The cenpB gene is not essential in mice. Chromosoma *107*, 570-576.

Keller, G. M. (1995). In vitro Differentiation of Embryonic Stem Cells. Curr. Opin. Cell Biol. *7*, 862-869.

Kellum, R. and Schedl, P. (1992). A group of scs elements function as domain boundaries in an enhancer-blocking assay. Mol. Cell. Biol. *12*, 2424-2431.

Keohane, A. M., O'Neill, L. P., Belyaev, N. D., Lavender, J. S., and Turner, B. M. (1996). X-Inactivation and histone H4 acetylation in embryonic stem cells. Dev. Biol. *180*, 618-30.

Kim, J., Sif, S., Jones, B., Jackson, A., Koipally, J., Heller, E., Winandy, S., Viel, A., Sawyer, A., Ikeda, T., Kingston, R., and Georgopoulos, K. (1999). Ikaros DNA-binding proteins direct formation of chromatin remodeling complexes in lymphocytes. Immunity. *10*, 345-355.

Kim, U. J., Han, M., Kayne, P., and Grunstein, M. (1988). Effects of histone H4 depletion on the cell cycle and transcription of Saccharomyces cerevisiae. EMBO J. *7*, 2211-2219.

Kimura, T., Mills, F. C., Allan, J., and Gould, H. (1983). Selective Unfolding Of Erythroid Chromatin In the Region Of the Active β-Globin Gene. Nature *306*, 709-712.

Kind, A. and Colman, A. (1999). Therapeutic cloning: Needs and prospects. Semin. Cell Dev. Biol. *10*, 279-286.

Kingston, R. E. and Narlikar, G. J. (1999). ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev. *13*, 2339-2352.

Kioussis, D. and Festenstein, R. (1997). Locus control regions: overcoming heterochromatin-induced gene inactivation in mammals. Curr. Opin. Genet. Dev. *7*, 614-619.

Kipling, D. (1995). The Telomere. (Oxford: Oxford University Press).

Kipling, D., Mitchell, A. R., Masumoto, H., Wilson, H. E., Nicol, L., and Cooke, H. J. (1995). CENP-B Binds a Novel Centromeric Sequence in the Asian Mouse Mus Caroli. Mol. Cell. Biol. *15*, 4009-4020.

Kipling, D., Wilson, H. E., Mitchell, A. R., Taylor, B. A., and Cooke, H. J. (1994). Mouse Centromere Mapping Using Oligonucleotide Probes That Detect Variants of the Minor Satellite. Chromosoma *103*, 46-55.

Kleinjan, D. J. and Van Heyningen, V. (1998). Position effect in human genetic disease. Hum. Mol. Genet. *7*, 1611-1618.

Klug, A., Rhodes, D., Smith, J., Finch, J. T., and Thomas, J. O. (1980). A low resolution structure for the histone core of the nucleosome. Nature *287*, 509-516.

Klug, M. G., Soonpaa, M. H., Koh, G. Y., and Field, L. J. (1996). Genetically selected cardiomyocytes from differentiating embronic stem cells form stable intracardiac grafts. J. Clin. Invest *98*, 216-224.

Kornberg, R. D. (1974). Chromatin structure: a repeating unit of histones and DNA. Science *184*, 868-871.

Kornberg, R. D. and Lorch, Y. (1991). Irresistible Force Meets Immovable Object: Transcription and the Nucleosome. Cell 833-836.

Kornberg, R. D. and Lorch, Y. (1992). Chromatin Structure and Transcription. Annu. Rev. Cell. Biol. *8*, 563-687.

Kornberg, R. D. and Lorch, Y. (1999). Twenty-five years of the nucleosome, fundamental particle of the eukaryote chromosome. Cell *98*, 285-294.

Kornberg, R. D. and Thomas, J. O. (1974). Chromatin structure; oligomers of the histones. Science *184*, 865-868.

Krajewski, W. A. and Becker, P. B. (1998). Reconstitution of hyperacetylated, DNase I-sensitive chromatin characterized by high conformational flexibility of nucleosomal DNA. Proc. Natl. Acad. Sci. U. S. A *95*, 1540-1545.

Krayev, A. S., Markusheva, T. V., Kramerov, D. A., Ryskov, A. P., Skryabin, K. G., Bayev, A. A., and Georgiev, G. P. (1982). Ubiquitous Transposon-Like Repeats B1 and B2 of the Mouse Genome - B2 Sequencing. Nucleic Acids Res. *10*, 7461-7475.

Krebs, J. E., Kuo, M. H., Allis, C. D., and Peterson, C. L. (1999). Cell cycle-regulated histone acetylation required for expression of the yeast HO gene. Genes Dev. *13*, 1412-1421.

Krude, T. and Knippers, R. (1991). Transfer of nucleosomes from parental to replicated chromatin. Mol. Cell Biol. *11*, 6257-6267.

Kruh, J. (1982). Effects of sodium butyrate, a new pharmacological agent, on cells in culture. Mol. Cell Biochem. *42*, 65-82.

Krumlauf, R., Holland, P. W., McVey, J. H., and Hogan, B. L. (1987). Developmental and spatial patterns of expression of the mouse homeobox gene, Hox 2.1. Development *99*, 603-617.

Kurz, A., Lampel, S., Nickolenko, J. E., Bradl, J., Benner, A., Zirbel, R. M., Cremer, T., and Lichter, P. (1996). Active and inactive genes localize preferentially in the periphery of chromosome territories. J. Cell Biol. *135*, 1195-1205.

Lambert, S., Muyldermans, S., Baldwin, J., Kilner, J., Ibel, K., and Wijns, L. (1991). Neutron scattering studies of chromatosomes. Biochem. Biophys. Res. Commun. *179*, 810-816.

Landsman, D. (1996). Histone H1 in Saccharomyces cerevisiae: a double mystery solved? Trends Biochem. Sci. *21*, 287-288.

Langmore, J. P. and Paulson, J. R. (1983). Low angle x-ray diffraction studies of chromatin structure in vivo and in isolated nuclei and metaphase chromosomes. J. Cell Biol. *96*, 1120-1131.

Langmore, J. P. and Schutt, C. (1980). The higher order structure of chicken erythrocyte chromosomes in vivo. Nature *288*, 620-622.

Lanini, L. and McKeon, F. (1995). Domains required for CENP-C assembly at the kinetochore. Mol. Biol. Cell *6*, 1049-1059.

Lee, D. Y., Hayes, J. J., Pruss, D., and Wolffe, A. P. (1993). A positive role for histone acetylation in transcription factor access to nucleosomal DNA. Cell *72*, 73-84.

Lee, H. L. and Archer, T. K. (1998). Prolonged glucocorticoid exposure dephosphorylates histone H1 and inactivates the MMTV promoter. EMBO J. *17*, 1454-1466.

Lee, H. W., Blasco, M. A., Gottlieb, G. J., Horner, J. W., Greider, C. W., and DePinho, R. A. (1998). Essential role of mouse telomerase in highly proliferative organs. Nature *392*, 569-574.

Lee, K. M. and Hayes, J. J. (1998). Linker DNA and H1-dependent reorganization of histone-DNA interactions within the nucleosome. Biochemistry *37*, 8622-8628.

Lennox, R. W. and Cohen, L. H. (1983). The histone H1 complements of dividing and nondividing cells of the mouse. J. Biol. Chem. *258*, 262-268.

Leuba, S. H., Bustamante, C., Van Holde, K. E., and Zlatanova, J. (1998a). Linker histone tails and N-tails of histone H3 are redundant: scanning force microscopy studies of reconstituted fibers. Biophys. J. *74*, 2830-2839.

Leuba, S. H., Bustamante, C., Zlatanova, J., and Van Holde, K. E. (1998b). Contributions of linker histones and histone H3 to chromatin structure: scanning force microscopy studies on trypsinized fibers. Biophys. J. *74*, 2823-2829.

Leuba, S. H., Yang, G., Robert, C., Samori, B., Van Holde, K. E., Zlatanova, J., and Bustamante, C. (1994). Three-dimensional structure of extended chromatin fibers as revealed by tapping-mode scanning force microscopy. Proc. Natl. Acad. Sci. U. S. A. *91*, 11621-11625.

Levy, A. and Noll, M. (1981). Chromatin fine structure of active and repressed genes. Nature *289*, 198-203.

Lewis, J. D., Meehan, R. R., Henzel, W. J., Maurerfogy, I., Jeppesen, P., Klein, F., and Bird, A. (1992). Purification, Sequence, and Cellular-Localization of a Novel Chromosomal Protein That Binds to Methylated DNA. Cell *69*, 905-914.

Li, M., Pevny, L., Lovell-Badge, R., and Smith, A. (1998). Generation of purified neural precursors from embryonic stem cells by lineage selection. Curr. Biol. *8*, 971-974.

Li, Y., Strahler, J. R., and Dodgson, J. B. (1997). Neither HMG-14a nor HMG-17 gene function is required for growth of chicken DT40 cells or maintenance of DNaseI-hypersensitive sites. Nucleic Acids Res. *25*, 283-288.

Lica, L. and Hamkalo, B. (1983). Preparation of centromeric heterochromatin by restriction endonuclease digestion of mouse L929 cells. Chromosoma *88*, 42-49.

Lin, S., Gaiano, N., Culp, P., Burns, J. C., Friedmann, T., Yee, J. K., and Hopkins, N. (1994). Integration and germ-line transmission of a pseudotyped retroviral vector in zebrafish. Science *265*, 666-669.

Linder, C. and Thoma, F. (1994). Histone H1 expressed in Saccharomyces cerevisiae binds to chromatin and affects survival, growth, transcription, and plasmid stability but does not change nucleosomal spacing. Mol. Cell Biol. *14*, 2822-2835.

Ling, X., Harkness, T. A. A., Schultz, M., Fisher-Adams, G., and Grunstein, M. (1996). Yeast histone H3 and H4 amino termini are important for nucleosome assembly in vivo and in vitro: redundant and position-independent functions in assembly but not in gene regulation. Genes Dev. *10*, 686-699.

Lingner, J. and Cech, T. R. (1998). Telomerase and chromosome end maintenance. Curr. Opin. Genet. Dev. *8*, 226-232.

Linney, E., Hardison, N. L., Lonze, B. E., Lyons, S., and DiNapoli, L. (1999). Transgene expression in zebrafish: A comparison of retroviral-vector and DNA-injection approaches. Dev. Biol. *213*, 207-216.

Linxweiler, W. and Horz, W. (1985). Reconstitution experiments show that sequence-specific histone-DNA interactions are the basis for nucleosome phasing on mouse satellite DNA. Cell *42*, 281-290.

Lipchitz, L. and Axel, R. (1976). Restriction endonuclease cleavage of satellite DNA in intact bovine nuclei. Cell *9*, 355-364.

Littlewood, T. D., Hancock, D. C., Danielian, P. S., Parker, M. G., and Evan, G. I. (1995). A modified oestrogen receptor ligand-binding domain as an improved switch for the regulation of heterologous proteins. Nucleic Acids Res. *23*, 1686-1690.

Loenig, U. E. (1969). The determination of the molecular weight of Ribonucleic acid by polyacrylamide-gel electrophoresis. Biochem. J. *113*, 131-138.

Lorch, Y., Zhang, M., and Kornberg, R. D. (1999). Histone octamer transfer by a chromatin-remodeling complex. Cell *96*, 389-392.

Lu, M. J., Mpoke, S. S., Dadd, C. A., and Allis, C. D. (1995). Phosphorylated and dephosphorylated linker histone H1 reside in distinct chromatin domains in Tetrahymena macronuclei. Mol. Biol. Cell *6*, 1077-1087.

Luger, K., Mader, A. W., Richmond, R. K., Sargent, D. F., and Richmond, T. J. (1997). Crystal structure of the nucleosome core particle at 2.8 A resolution. Nature *389*, 251-260.

Lupton, S. D., Brunton, L. L., Kalberg, V. A., and Overell, R. W. (1991). Dominant positive and negative selection using a hygromycin phosphotransferase-thymidine kinase fusion gene. Mol. Cell Biol. *11*, 3374-3378.

Lyon, M. F. (1999). X-chromosome inactivation. Curr. Biol. *9*, R235-R237.

Mackay, A. M., Ainsztein, A. M., Eckley, D. M., and Earnshaw, W. C. (1998). A dominant mutant of inner centromere protein (INCENP), a chromosomal protein, disrupts prometaphase congression and cytokinesis. J. Cell Biol. *140*, 991-1002.

Madisen, L., Krumm, A., Hebbes, T. R., and Groudine, M. (1998). The immunoglobulin heavy chain locus control region increases histone acetylation along linked c-myc genes. Mol. Cell Biol. *18*, 6281-6292.

Makarov, V. L., Lejnine, S., Bedoyan, J., and Langmore, J. P. (1993). Nucleosomal organization of telomere-specific chromatin in rat. Cell *73*, 775-787.

Manders, E. M., Kimura, H., and Cook, P. R. (1999). Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. J. Cell Biol. *144*, 813-821.

Manuelidis, L. (1981). Consensus Sequence of Mouse Satellite DNA Indicates It Is Derived From Tandem 116 Basepair Repeats. FEBS Lett. *129*, 25-28.

Manuelidis, L. (1990). A View of Interphase Chromosomes. Science *250*, 1533-1540.

Manuelidis, L. and Chen, T. L. (1990). A unified model of eukaryotic chromosomes. Cytometry *11*, 8-25.

Mao, X., Fujiwara, Y., and Orkin, S. H. (1999). Improved reporter strain for monitoring Cre recombinase-mediated DNA excisions in mice. Proc. Natl. Acad. Sci. U. S. A. *96*, 5037-5042.

Marion, C., Roche, J., Roux, B., and Gorka, C. (1985). Differences in the condensation of chromatin by individual subfractions of histone H1: implications for the role of H1° in the structural organization of chromatin. Biochemistry *24*, 6328-6335.

Marsden, M. P. and Laemmli, U. K. (1979). Metaphase chromosome structure: evidence for a radial loop model. Cell *17*, 849-858.

Martin, G. R. (1980). Teratocarcinomas and Mammalian Embryogenesis. Science *209*, 768-776.

Martin, G. R. (1981). Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. Proc. Natl. Acad. Sci. U. S. A. *78*, 7634-7638.

Masumoto, H., Ikeno, M., Nakano, M., Okazaki, T., Grimes, B., Cooke, H., and Suzuki, N. (1998). Assay of centromere function using a human artificial chromosome. Chromosoma *107*, 406-416.

Masumoto, H., Masukata, H., Muro, Y., Nozaki, N., and Okazaki, T. (1989). A human centromere antigen (CENP-B) interacts with a short specific sequence in alphoid DNA, a human centromeric satellite. J. Cell Biol. *109*, 1963-1973.

Mathew, C. G. P., Goodwin, G. H., Igokemenes, T., and Johns, E. W. (1981). The Protein-Composition of Rat Satellite Chromatin. FEBS Lett. *125*, 25-29.

Matsuda, T., Nakamura, T., Nakao, K., Arai, T., Katsuki, M., Heike, T., and Yokota, T. (1999). STAT3 activation is sufficient to maintain an undifferentiated state of mouse embryonic stem cells. EMBO J. *18*, 4261-4269.

Mazen, A., Hacques, M. F., and Marion, C. (1987). H3 phosphorylation-dependent structural changes in chromatin. Implications for the role of very lysine-rich histones. J. Mol. Biol. *194*, 741-745.

Mazrimas, J. A., Balhorn, R., and Hatch, F. T. (1979). Separation of satellite DNA chromatin and main band DNA chromatin from mouse brain. Nucleic Acids Res. *7*, 935-46.

McCall, K. and Bender, W. (1996). Probes of chromatin accessibility in the Drosophila bithorax complex respond differently to Polycomb-mediated repression. EMBO J. *15*, 569-580.

McGhee, J. D., Nickol, J. M., Felsenfeld, G., and Rau, D. C. (1983b). Higher Order Structure of Chromatin: Orientation of Nucleosomes within the 30nm Chromatin Solenoid is Independant of Species and Spacer Length. Cell *33*, 831-841.

McGhee, J. D., Nickol, J. M., Felsenfeld, G., and Rau, D. C. (1983a). Histone hyperacetylation has little effect on the higher order folding of chromatin. Nucleic Acids Res. *11*, 4065-4075.

McKay, D. B. and Steitz, T. A. (1981). Structure of catabolite gene activator protein at 2.9 A resolution suggests binding to left-handed B-DNA. Nature *290*, 744-749.

Meehan, R. R., Lewis, J. D., McKay, S., Kleiner, E. L., and Bird, A. P. (1989). Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell *58*, 499-507.

Meersseman, G., Pennings, S., and Bradbury, E. M. (1992). Mobile nucleosomes--a general behavior. EMBO J. *11*, 2951-2959.

Meluh, P. B. and Koshland, D. (1997). Budding yeast centromere composition and assembly as revealed by in vivo cross-linking. Genes Dev. *11*, 3401-3412.

Meluh, P. B., Yang, P., Glowczewski, L., Koshland, D., and Smith, M. M. (1998). Cse4p is a component of the core centromere of Saccharomyces cerevisiae. Cell *94*, 607-613.

Mendelson, E. and Bustin, M. (1984). Monoclonal-Antibodies Against Distinct Determinants Of Histone-H5 Bind to Chromatin. Biochemistry *23*, 3459-3466.

Michelotti, E. F., Sanford, S., and Levens, D. (1997). Marking of active genes on mitotic chromosomes. Nature *388*, 895-899.

Miklos, G. L. and John, B. (1979). Heterochromatin and satellite DNA in man: properties and prospects. Am. J. Hum. Genet. *31* , 264-80.

Mileham, P. and Brown, S. D. M. (1994). The pufferfish genome: Small is beautiful. Bioessays *16*, 153-154.

Mills, W., Critcher, R., Lee, C., and Farr, C. J. (1999). Generation of an approximately 2.4 Mb human X centromere-based minichromosome by targeted telomere-associated chromosome fragmentation in DT40. Hum. Mol. Genet. *8*, 751-761.

Misteli, T., Caceres, J. F., and Spector, D. L. (1997). The dynamics of a pre-mRNA splicing factor in living cells. Nature *387*, 523-527.

Misteli, T. and Spector, D. L. (1998). The cellular organization of gene expression. Curr. Opin. Cell Biol. *10*, 323-331.

Mitchell, A. R., Gosden, J. R., and Miller, D. A. (1985). A cloned sequence, p82H, of the alphoid repeated DNA family found at the centromeres of all human chromosomes. Chromosoma *92*, 369-77.

Mitchell, A. R., Jeppesen, P., Nicol, L., Morrison, H., and Kipling, D. (1996). Epigenetic control of mammalian centromere protein binding: Does DNA methylation have a role? J. Cell Sci. *109*, 2199-2206.

Mizzen, C. A., Dou, Y., Liu, Y., Cook, R. G., Gorovsky, M. A., and Allis, C. D. (1999). Identification and mutation of phosphorylation sites in a linker histone. Phosphorylation of macronuclear H1 is not essential for viability in tetrahymena. J. Biol. Chem. *274*, 14533-14536.

Moroi, Y., Peebles, C., Fritzler, M. J., Steigerwald, J., and Tan, E. M. (1980). Autoantibody to centromere (kinetochore) in scleroderma sera. Proc. Natl. Acad. Sci. U. S. A. *77*, 1627-1631.

Mountford, P., Zevnik, B., Duwel, A., Nichols, J., Li, M., Dani, C., Robertson, M., Chambers, I., and Smith, A. (1994). Dicistronic targeting constructs: Reporters and modifiers of mammalian expression. Proc. Natl. Acad. Sci. U. S. A. *91*, 4303-4307.

Muller, H. J. (1938). The remaking of chromosomes. Collect. Net *8*, 182-195.

Mura, C. V., Setterfield, G., and Neelin, J. M. (1982). Immunofluorescent Localization Of Lysine-Rich Histones In Isolated- Nuclei From Adult and Embryonic Chicken Erythrocytes. Can. J. Biochem. *60*, 215-223.

Murray, A. W. and Szostak, J. W. (1983). Construction of artificial chromosomes in yeast. Nature *305*, 189-193.

Musich, P. R., Brown, F. L., and Maio, J. J. (1977). Subunit structure of chromatin and the organization of eukaryotic highly repetitive DNA: Nucleosomal proteins associated with a hihgly repetitive mammalian DNA. Proc. Natl. Acad. Sci. U. S. A. *74*, 3297-3301.

Mutskov, V., Gerber, D., Angelov, D., Ausio, J., Workman, J., and Dimitrov, S. (1998). Persistent interactions of core histone tails with nucleosomal DNA following acetylation and transcription factor binding. Mol. Cell Biol. *18*, 6293-6304.

Nagy, A., Gocza, E., Diaz, E. M., Prideaux, V. R., Ivanyi, E., Markkula, M., and Rossant, J. (1990). Embryonic Stem-Cells Alone Are Able to Support Fetal Development in the Mouse. Development *110*, 815.

Nagy, A., Rossant, J., Nagy, R., Abramownewerly, W., and Roder, J. C. (1993). Derivation Of Completely Cell Culture-Derived Mice From Early-Passage Embryonic Stem-Cells. Proc. Natl. Acad. Sci. U. S. A. *90* , 8424-8428.

Nakamura, T. M., Morin, G. B., Chapman, K. B., Weinrich, S. L., Andrews, W. H., Lingner, J., Harley, C. B., and Cech, T. R. (1997). Telomerase catalytic subunit homologs from fission yeast and human. Science *277*, 955-959.

Nakano, T., Kodama, H., and Honjo, T. (1996). In vitro development of primitive and definitive erythrocytes from different precursors. Science *272*, 722-724.

Nan, X., Cross, S., and Bird, A. (1998b). Gene silencing by methyl-CpG-binding proteins. Novartis. Found. Symp. *214*, 6-16.

Nan, X., Ng, H. H., Johnson, C. A., Laherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998a). Transcriptional repression by the methyl-CpG-binding protein MeCP2 involves a histone deacetylase complex. Nature *393*, 386-389.

Nichols, J., Evans, E. P., and Smith, A. G. (1990). Establishment of germ-line-competent embryonic stem (ES) cells using differentiation inhibiting activity. Development *110*, 1341-1348.

Nichols, J., Zevnik, B., Anastassiadis, K., Niwa, H., Klewe-Nebenius, D., Chambers, I., Scholer, H., and Smith, A. (1998). Formation of Pluripotent Stem Cells in the Mammalian Embryo Depends on the POU Transcription Factor Oct-4. Cell *95*, 379-391.

Niwa, H., Burdon, T., Chambers, I., and Smith, A. (1998). Self-renewal of pluripotent embryonic stem cells is mediated via activation of STAT3. Genes Dev. *12*, 2048-2060.

Niwa, H., Yamamura, K., and Miyazaki, J. (1991). Efficient Selection For High-Expression Transfectants With a Novel Eukaryotic Vector. Gene *108*, 193-199.

Noll, M. (1974). Subunit Structure of Chromatin. Nature *254*, 249-251.

Norton, V. G., Imai, B. S., Yau, P., and Bradbury, E. M. (1989). Histone acetylation reduces nucleosome core particle linking number change. Cell *57*, 449-457.

Notbohm, H. (1986). Small angle scattering of cell nuclei. Eur. Biophys. J. *13*, 367-372.

Nugent, C. I. and Lundblad, V. (1998). The telomerase reverse transcriptase: components and regulation. Genes Dev. *12*, 1073-1085.

O'Brien, I. E., Reutelingsperger, C. P., and Holdaway, K. M. (1997). Annexin-V and TUNEL use in monitoring the progression of apoptosis in plants. Cytometry *29*, 28-33.

O'Neill, L. P. and Turner, B. M. (1995). Histone H4 Acetylation Distinguishes Coding Regions Of the Human Genome From Heterochromatin In a Differentiation-Dependent But Transcription-Independent Manner. EMBO J. *14*, 3946-3957.

Oberto, J., Drlica, K., and Rouviere-Yaniv, J. (1994). Histones, HMG, HU, IHF: Meme combat. Biochimie *76*, 901-908.

Ogawa, M. (1993). Differentiation and proliferation of hematopoietic stem cells. Blood *81*, 2844-2853.

Ohsumi, K., Katagiri, C., and Kishimoto, T. (1993). Chromosome condensation in Xenopus mitotic extracts without histone H1. Science *262*, 2033-2035.

Omori, A., Igo-Kemenes, T., and Zachau, H. G. (1980). Different repeat lengths in rat satellite I DNA containing chromatin and bulk chromatin. Nucleic Acids Res. *8*, 5363-75.

Ortiz, B. D., Cado, D., and Winoto, A. (1999). A new element within the T-cell receptor alpha locus required for tissue-specific locus control region activity. Mol. Cell Biol. *19*, 1901-1909.

Osborne, H. B. and Chabanas, A. (1984). Kinetics Of Histone H-1 Accumulation and Commitment to Differentiation In Murine Erythroleukemia-Cells. Exp. Cell Res. *152*, 449-458.

Oudet, P., Gross-Bellard, M., and Chambon, P. (1975). Electron microscopic and biochemical evidence that chromatin structure is a repeating unit. Cell *4*, 281-300.

Page, S. L., Earnshaw, W. C., Choo, K. H., and Shaffer, L. G. (1995). Further evidence that CENP-C is a necessary component of active centromeres: studies of a dic(X; 15) with simultaneous immunofluorescence and FISH. Hum. Mol. Genet. *4*, 289-294.

Panetta, G., Buttinelli, M., Flaus, A., Richmond, T. J., and Rhodes, D. (1998). Differential nucleosome positioning on Xenopus oocyte and somatic 5 S RNA genes determines both TFIIIA and H1 binding: a mechanism for selective H1 repression. J. Mol. Biol. *282*, 683-697.

Pardue, M. L. and DeBaryshe, P. G. (1999). Telomeres and telomerase: more than the end of the line. Chromosoma *108*, 73-82.

Pardue, M. L. and Gall, J. G. (1970). Chromosomal localization of mouse satellite DNA. Science *168*, 1356-1358.

Paro, R. (1990). Imprinting a determined state into the chromatin of Drosophila. Trends Genet. *6*, 416-421.

Pashev, I. G., Dimitrov, S. I., Ivanov, I. G., and Markov, G. G. (1983). Histone Acetylation In Chromatin Containing Mouse Satellite DNA. Eur. J. Biochem. *133*, 379-382.

Patterton, H. G., Landel, C. C., Landsman, D., Peterson, C. L., and Simpson, R. T. (1998). The biochemical and phenotypic characterization of Hho1p, the putative linker histone H1 of Saccharomyces cerevisiae. J. Biol. Chem. *273*, 7268-7276.

Paulson, J. R. and Laemmli, U. K. (1977). The structure of histone-depleted metaphase chromosomes. Cell *12*, 817-828.

Paulson, J. R. and Langmore, J. P. (1983). Low angle x-ray diffraction studies of HeLa metaphase chromosomes: effects of histone phosphorylation and chromosome isolation procedure. J. Cell Biol. *96*, 1132-1137.

Pearson, E. C., Bates, D. L., Prospero, T. D., and Thomas, J. O. (1984). Neuronal nuclei and glial nuclei from mammalian cerebral cortex. Nucleosome repeat lengths, DNA contents and H1 contents. Eur. J. Biochem. *144*, 353-360.

Pehrson, J. and Cole, R. D. (1980). Histone H1° Accumulates in Growth-inhibited Cultured Cells. Nature *285*, 43-44.

Penman, S. (1995). Rethinking Cell Structure. Proc. Natl. Acad. Sci. U. S. A. *92*, 5251-5257.

Pereira, S. L. and Reeve, J. N. (1998). Histones and nucleosomes in Archaea and Eukarya: a comparative analysis. Extremophiles. *2*, 141-148.

Perez-Castro, A. V., Shamanski, F. L., Meneses, J. J., Lovato, T. L., Vogel, K. G., Moyzis, R. K., and Pedersen, R. (1998). Centromeric protein B null mice are viable with no apparent abnormalities. Dev. Biol. *201*, 135-143.

Perez-Martin, J. and Johnson, A. D. (1998). The C-terminal domain of Sin1 interacts with the SWI-SNF complex in yeast. Mol. Cell Biol. *18*, 4157-4164.

Perry, C. A. and Annunziato, A. T. (1989). Influence of histone acetylation on the solubility, H1 content and DNase I sensitivity of newly assembled chromatin. Nucleic Acids Res. *17*, 4275-4291.

Pestana, A. and Pitot, H. C. (1974). N-terminal acetylation of histone-like nascent peptides on rat liver polyribosomes in vitro. Nature *247*, 200-202.

Peterson, C. L. and Tamkun, J. W. (1995). The SWI-SNF complex: a chromatin remodeling machine? Trends Biochem. Sci. *20*, 143-146.

Peterson, C. L., Zhao, Y. M., and Chait, B. T. (1998). Subunits of the yeast SWI/SNF complex are members of the actin- related protein (ARP) family. J. Biol. Chem. *273*, 23641-23644.

Pienta, K. J. and Coffey, D. S. (1984a). A structural analysis of the role of the nuclear matrix and DNA loops in the organization of the nucleus and chromosome. J. Cell Sci. Suppl *1*, 123-135.

Pietras, D. F., Bennett, K. L., Siracusa, L. D., Woodworthgutai, M., Chapman, V. M., Gross, K. W., Kanehaas, C., and Hastie, N. D. (1983). Construction of a Small Mus-Musculus Repetitive Dna Library - Identification of a New Satellite Sequence in Mus-Musculus. Nucleic Acids Res. *11*, 6965-6983.

Pirrotta, V. (1998). Polycombing the genome: PcG, trxG, and chromatin silencing. Cell *93*, 333-336.

Pirrotta, V. and Rastelli, L. (1994). White gene expression, repressive chromatin domains and homeotic gene regulation in Drosophila. Bioessays *16*, 549-556.

Pluta, A. F., Mackay, A. M., Ainsztein, A. M., Goldberg, I. G., and Earnshaw, W. C. (1995). The Centromere - Hub of Chromosomal Activities. Science *270*, 1591-1594.

Pluta, A. F., Saitoh, N., Goldberg, I., and Earnshaw, W. C. (1992). Identification of a Subdomain of CENP-B That is Necessary and Sufficient for Localization to the Human Centromere. J. Cell Biol. *116*, 1081-1093.

Pluthero, F. G. (1993). Rapid purification of high-activity Taq DNA polymerase. Nucleic Acids Res. *21*, 4850-4851.

Pollard, K. J. and Peterson, C. L. (1997). Role for ADA/GCN5 products in antagonizing chromatin-mediated transcriptional repression. Mol. Cell Biol. *17*, 6212-6222.

Pombo, A., Jackson, D. A., Hollinshead, M., Wang, Z. X., Roeder, R. G., and Cook, P. R. (1999). Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. EMBO J. *18*, 2241-2253.

Postnikov, Y. V., Herrera, J. E., Hock, R., Scheer, U., and Bustin, M. (1997). Clusters of nucleosomes containing chromosomal protein HMG-17 in chromatin. J. Mol. Biol. *274*, 454-465.

Prosser, J., Frommer, M., Paul, C., and Vincent, P. C. (1986). Sequence relationships of three human satellite DNAs. J. Mol. Biol. *187*, 145-155.

Pruss, D., Bartholomew, B., Persinger, J., Hayes, J., Arents, G., Moudrianakis, E. N., and Wolffe, A. P. (1996). An asymmetric model for the nucleosome: A binding site for linker histones inside the DNA gyres. Science *274*, 614-617.

Pruss, D., Hayes, J. J., and Wolffe, A. P. (1995). Nucleosomal anatomy - where are the histones? Bioessays *17*, 161-170.

Radic, M. Z., Lundgren, K., and Hamkalo, B. A. (1987). Curvature of Mouse Satellite DNA and Condensation of Heterochromatin. Cell *50*, 1101-1108.

Radic, M. Z., Saghbini, M., Elton, T. S., Reeves, R., and Hamkalo, B. A. (1992). Hoechst-33258, Distamycin-a, and High Mobility Group Protein-I (HMG- I) Compete For Binding to Mouse Satellite DNA. Chromosoma *101*, 602-608.

Ramakrishnan, V. (1997). Histone H1 and chromatin higher-order structure. Crit. Rev. Eukaryot. Gene Express. *7*, 215-230.

Ramakrishnan, V., Finch, J. T., Graziano, V., Lee, P. L., and Sweet, R. M. (1993). Crystal-Structure of Globular Domain of Histone H5 and Its Implications For Nucleosome Binding. Nature *362*, 219-223.

Rattner, J. B. and Lin, C. C. (1985). Radial loops and helical coils coexist in metaphase chromosomes. Cell *42*, 291-296.

Raz, R., Lee, C. K., Cannizzaro, L. A., D'Eustachio, p., and Levy, D. E. (1999). Essential role of STAT3 for embryonic stem cell pluripotency. Proc. Natl. Acad. Sci. U. S. A. *96*, 2846-2851.

Recht, J. and Osley, M. A. (1999). Mutations in both the structured domain and N-terminus of histone H2B bypass the requirement for Swi-Snf in yeast. EMBO J. *18*, 229-240.

Reeves, R. and Nissen, M. S. (1990). The A.T-DNA-Binding Domain Of Mammalian High Mobility Group-I Chromosomal-Proteins - a Novel Peptide Motif For Recognizing DNA- Structure. J. Biol. Chem. *265*, 8573-8582.

Remboutsika, E., Lutz, Y., Gansmuller, A., Vonesch, J. L., Losson, R., and Chambon, P. (1999). The putative nuclear receptor mediator TIF1 α is tightly associated with euchromatin. J. Cell Sci. *112 (Pt 11)*, 1671-1683.

Renz, M., Nehls, P., and Hozier, J. (1977). Involvement of histone H1 in the organization of the chromosome fiber. Proc. Natl. Acad. Sci. U. S. A. *74*, 1879-1883.

Richmond, T. J., Finch, J. T., Rushton, B., Rohdes, D., and Klug, A. (1984). Structure of the nucleosome core particle at 7A resolution. Nature *311*, 532-537.

Riggs, M. G., Whittaker, R. G., Neumann, J. R., and Ingram, V. M. (1977). n-Butyrate causes histone modification in HeLa and Friend erythroleukaemia cells. Nature *268*, 462-464.

Robertson, E., Bradley, A., Kuehn, M., and Evans, M. (1986). Germ-line transmission of genes introduced into cultured pluripotential cells by retroviral vector. Nature *323*, 445-448.

Robertson, E. J. (1986). Pluripotential stem-cell lines as a route into the mouse germ line. Trends Genet. *2*, 9-13.

Roche, J., Gorka, C., Goeltz, P., and Lawrence, J. J. (1985). Association Of Histone H1° With a Gene Repressed During Liver Development. Nature *314*, 197-198.

Rogakou, E. P., Boon, C., Redon, C., and Bonner, W. M. (1999). Megabase chromatin domains involved in DNA double-strand breaks In vivo. J. Cell Biol. *146*, 905-916.

Romano, G. (1992). Histone variants during sea urchin development. Cell Biol. Int. Rep. *16*, 197-206.

Romanova, L. Y., Deriagin, G. V., Mashkova, T. D., Tumeneva, I. G., Mushegian, A. R., Kisselev, L. L., and Alexandrov, I. A. (1996). Evidence for selection in evolution of α-satellite DNA: the central role of CENP-B/pJα binding region. J. Mol. Biol. *261*, 334-340.

Rose, S. M. and Garrard, W. T. (1984). Differentiation-dependent chromatin alterations precede and accompany transcription of immunoglobulin light chain genes. J. Biol. Chem. *259*, 8534-8544.

Rosner, M. H., Vigano, M. A., Ozato, K., Timmons, P. M., Poirier, F., Rigby, P. W., and Staudt, L. M. (1990). A POU-domain transcription factor in early stem cells and germ cells of the mammalian embryo. Nature *345*, 686-692.

Rousseau, D., Khochbin, S., Gorka, C., and Lawrence, J.-J. (1991). Regulation of Histone H1° Accumulation During Induced Differentiation of Murine Erythroleukemia Cells. J. Mol. Biol. *217*, 85-92.

Rousseau, S., Asselin, M., Renaud, J., and Ruizcarrillo, A. (1993). Transcription Of the Histone H5-Gene Is Regulated By 3 Differentiation-Specific Enhancers. Mol. Cell. Biol. *13*, 4904-4917.

Rudolph, K. L., Chang, S., Lee, H. W., Blasco, M., Gottlieb, G. J., Greider, C., and DePinho, R. A. (1999). Longevity, stress response, and cancer in aging telomerase-deficient mice. Cell *96*, 701-712.

Ruiz-Carrillo, A., Wangh, L. J., and Allfrey, V. G. (1976). Selective Synthesis and Modification of Nuclear Proteins during Maturation of Avian Erythroid Cells. Arch. Biochem. Biophys. *174*, 273-290.

Ruiz-Carrillo, A., Wangh, L. J., Littau, V. C., and Allfrey, V. G. (1974). Changes in Histone Acetyl Content and in Nuclear Non-Histone Protein Composition of Avian Erythroid Cells at Different Stages of Maturation. J. Biol. Chem. *249*, 7358-7368.

Rundlett, S. E., Carmen, A. A., Kobayashi, R., Bavykin, S., Turner, B. M., and Grunstein, M. (1996). HDA1 and RPD3 are members of distinct yeast histone deacetylase complexes that regulate silencing and transcription. Proc. Natl. Acad. Sci. U. S. A. *93*, 14503-14508.

Rydberg, B., Holley, W. R., Mian, I. S., and Chatterjee, A. (1998). Chromatin conformation in living cells: support for a zig-zag model of the 30 nm chromatin fiber. J. Mol. Biol. *284*, 71-84.

Sabapathy, K., Klemm, M., Jaenisch, R., and Wagner, E. F. (1997). Regulation of ES cell differentiation by functional and conformational modulation of p53. EMBO J. *16*, 6217-6229.

Sachs, R. K., van den Engh, G., Trask, B., Yokota, H., and Hearst, J. E. (1995). A random-walk/giant-loop model for interphase chromosomes. Proc. Natl. Acad. Sci. U. S. A. *92*, 2710-2714.

Sahasrabuddhe, C. G. and Van Holde, K. E. (1974). The effect of trypsin on nuclease-resistant chromatin fragments. J. Biol. Chem. *249*, 152-156.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). Molecular cloning: A laboratory manual. (New York: Coldspring Harbor Press).

Sandell, L. L. and Zakian, V. A. (1993). Loss of a yeast telomere: arrest, recovery, and chromosome loss. Cell *75*, 729-739.

Sandman, K., Pereira, S. L., and Reeve, J. N. (1998). Diversity of prokaryotic chromosomal proteins and the origin of the nucleosome. Cell Mol. Life Sci. *54*, 1350-1364.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. U. S. A. *74*, 5463-5467.

Sauer, B. and Henderson, N. (1990). Targeted insertion of exogenous DNA into the eukaryotic genome by the Cre recombinase. New Biol. *2*, 441-9.

Sauve, D. M., Anderson, H. J., Ray, J. M., James, W. M., and Roberge, M. (1999). Phosphorylation-induced rearrangement of the histone H3 NH2-terminal domain during mitotic chromosome condensation. J. Cell Biol. *145*, 225-235.

Scholer, H. R., Ruppert, S., Suzuki, N., Chowdhury, K., and Gruss, P. (1990). New type of POU domain in germ line-specific protein Oct-4. Nature *344*, 435-439.

Schroth, G. P., Yau, P., Imai, B. S., Gatewood, J. M., and Bradbury, E. M. (1990). A NMR study of mobility in the histone octamer. FEBS Lett. *268*, 117-120.

Schultz, S. C., Shields, G. C., and Steitz, T. A. (1991). Crystal structure of a CAP-DNA complex: the DNA is bent by 90 degrees. Science *253*, 1001-1007.

Selig, S., Ariel, M., Goitein, R., Marcus, M., and Cedar, H. (1988). Regulation Of Mouse Satellite Dna-Replication Time. EMBO J. *7*, 419-426.

Shamblott, M. J., Axelman, J., Wang, S., Bugg, E. M., Littlefield, J. W., Donovan, P. J., Blumenthal, P. D., Huggins, G. R., and Gearhart, J. D. (1998). Derivation of pluripotent stem cells from cultured human primordial germ cells. Proc. Natl. Acad. Sci. U. S. A *95*, 13726-13731.

Shao, Z., Raible, F., Mollaaghababa, R., Guyon, J. R., Wu, C. T., Bender, W., and Kingston, R. E. (1999). Stabilization of chromatin structure by PRC1, a Polycomb complex. Cell *98*, 37-46.

Shelby, R. D., Vafa, O., and Sullivan, K. F. (1997). Assembly of CENP-A into centromeric chromatin requires a cooperative array of nucleosomal DNA contact sites. J. Cell Biol. *136*, 501-513.

Shen, C.-H. (1997) The role of linker histone globular domains in chromatosome formation. Ph.D Thesis, The University of Edinburgh

Shen, X. and Gorovsky, M. A. (1996). Linker histone H1 regulates specific gene expression but not global transcription in vivo. Cell *86*, 475-83.

Shen, X., Yu, L., Weir, J. W., and Gorovsky, M. A. (1995). Linker histones are not essential and affect chromatin condensation in vivo. Cell *82*, 47-56.

Sherman, M. I. and Miller, R. A. (1978). F9 Embryonal Carcinoma Cells Can Differentiate into Endoderm-Like Cells. Dev. Biol. *63*, 27-34.

Shibahara, K. and Stillman, B. (1999). Replication-dependent marking of DNA by PCNA facilitates CAF-1-coupled inheritance of chromatin. Cell *96*, 575-585.

Shibata, K., Inagaki, M., and Ajiro, K. (1990). Mitosis-specific histone H3 phosphorylation in vitro in nucleosome structures. Eur. J. Biochem. *192*, 87-93.

Shiels, P. G., Kind, A. J., Campbell, K. H., Waddington, D., Wilmut, I., Colman, A., and Schnieke, A. E. (1999). Analysis of telomere lengths in cloned sheep. Nature *399*, 316-317.

Simpson, R. T. (1978b). Structure of chromatin containing extensively acetylated H3 and H4. Cell *13*, 691-699.

Simpson, R. T. (1978a). Structure of the chromatosome, a chromatin particle containing 160 base pairs of DNA and all the histones. Biochemistry *17*, 5524-31.

Simpson, R. T. (1991). Nucleosome positioning: occurrence, mechanisms, and functional consequences. Progress in Nucleic Acid Research and Molecular Biology *40*, 143-184.

Simpson, R. T., Whitlock, J. P. J., Bina-Stein, M., and Stein, A. (1978). Histone-DNA interactions in chromatin core particles. Cold Spring Harb. Symp. Quant. Biol. *42 Pt 1:127-36*, 127-136.

Singer, M. F. (1982). Highly Repeated Sequences in Mammalian Genomes. Int. Rev. Cytol. *76*, 67-112.

Sirotkin, A. M., Edelmann, W., Cheng, G. H., Kleinszanto, A., Kucherlapati, R., and Skoultchi, A. I. (1995). Mice Develop Normally Without the H1° Linker Histone. Proc. Natl. Acad. Sci. U. S. A. *92*, 6434-6438.

Sleigh, M. J. (1992). Differentiation and Proliferation In Mouse Embryonal Carcinoma-Cells. Bioessays *14*, 769-775.

Smith, A. G. (1991). Culture and Differentiation of Embryonic Stem Cells. J. Tissue Culture Methods *13*, 89-94.

Smith, A. G., Heath, J. K., Donaldson, D. D., Wong, G. G., Moreau, J., Stahl, M., and Rogers, D. (1988). Inhibition of Pluripotential Embryonic Stem-Cell Differentiation By Purified Polypeptides. Nature *336*, 688-690.

Smith, A. J. H., De Sousa, M. A., Kwabi-Addo, G., Heppell-Parton, A., Impey, H., and Rabbitts, P. (1995). A site-directed chromosomal translocation induced in embryonic stem cells by Cre-loxP recombination. Nat. Genet. *9*, 376-385.

Smith, P. A., Jackson, V., and Chalkley, R. (1984). Two-stage maturation process for newly replicated chromatin. Biochemistry *23*, 1576-1581.

Sobel, R. E., Cook, R. G., Perry, C. A., Annunziato, A. T., and Allis, C. D. (1995). Conservation of deposition-related acetylation sites in newly synthesized histones H3 and H4. Proc. Natl. Acad. Sci. U. S. A. *92*, 1237-1241.

Solomon, M. J., Strauss, F., and Varshavsky, A. (1986). A Mammalian High Mobility Group Protein Recognizes Any Stretch Of 6 A.T Base-Pairs In Duplex DNA. Proc. Natl. Acad. Sci. U. S. A. *83*, 1276-1280.

Solter, D., Shevinsky, L., Knowles, B. B., and Strickland, S. (1979). The Induction of Antigenic Changes in a Teratocarcinoma Stem Cell Line (F9) by Retinoic Acid. Dev. Biol. *70*, 515-521.

Southern, E. M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. *98*, 503-517.

Spadafora, C., Bellard, M., Compton, J. L., and Chambon, P. (1976). The DNA repeat lengths in chromatins from sea urchin sperm and gastrule cells are markedly different. FEBS Lett. *69*, 281-285.

Starr, D. A., Williams, B. C., Li, Z., Etemad-Moghadam, B., Dawe, R. K., and Goldberg, M. L. (1997). Conservation of the centromere/kinetochore protein ZW10. J. Cell Biol. *138*, 1289-1301.

Staynov, D. Z. and Crane-Robinson, C. (1988). Footprinting of Linker Histones-H5 and Histones-H1 On the Nucleosome. EMBO J. *7*, 3685-3691.

Stedman, E. and Stedman, E. (1950). Cell Specificity of Histones. Nature *166*, 780-781.

Steger, D. J., Eberharter, A., John, S., Grant, P. A., and Workman, J. L. (1998). Purified histone acetyltransferase complexes stimulate HIV-1 transcription from preassembled nucleosomal arrays. Proc. Natl. Acad. Sci. U. S. A. *95*, 12924-12929.

Stein, A. and Mitchell, M. (1988). Generation of different nucleosome spacing periodicities in vitro. Possible origin of cell type specificity. J. Mol. Biol. *203*, 1029-1043.

Steinbach, O. C., Wolffe, A. P., and Rupp, R. A. W. (1997). Somatic linker histones cause loss of mesodermal competence in Xenopus. Nature *389*, 395-399.

Strausbaugh, L. D. and Williams, S. M. (1996). High density of an SAR-associated motif differentiates heterochromatin from euchromatin. J. Theor. Biol. *183*, 159-167.

Strauss, F. and Varshavsky, A. (1984). A Protein Binds to a Satellite DNA Repeat At 3 Specific Sites That Would Be Brought Into Mutual Proximity By DNA Folding In the Nucleosome. Cell *37*, 889-901.

Strick, R. and Laemmli, U. K. (1995). SARs Are cis DNA Elements Of Chromosome Dynamics - Synthesis Of a SAR Repressor Protein. Cell *83*, 1137-1148.

Strickland, S. and Mahdavi, V. (1978). The Induction of Differentiation in Teratocarcinoma Stem Cells by Retinoic Acid. Cell *15*, 393-403.

Strickland, S., Smith, K. K., and Marotti, K. R. (1980). Hormomal Induction of Differentiation in Teratocarcinoma Stem Cells: Generation of Parietal Endoderm by Retinoic Acid and Dibutyryl cAMP. Cell *21*, 347-355.

Strouboulis, J. and Wolffe, A. P. (1996). Functional Compartmentalization Of the Nucleus. J. Cell Sci. *109*, 1991-2000.

Struhl, K. (1998). Histone acetylation and transcriptional regulatory mechanisms. Genes Dev. *12*, 599-606.

Struhl, K. (1999). Fundamentally different logic of gene regulation in eukaryotes and prokaryotes. Cell *98*, 1-4.

Sullivan, B. A. and Schwartz, S. (1995). Identification of centromeric antigens in dicentric Robertsonian translocations: CENP-C and CENP-E are necessary components of functional centromeres. Hum. Mol. Genet. *4*, 2189-2197.

Sullivan, K. F., Hechenberger, M., and Masri, K. (1994). Human CENP-A Contains a Histone H3 Related Histone Fold Domain That Is Required For Targeting to the Centromere. J. Cell Biol. *127*, 581-592.

Sun, J. M., Ali, Z., Lurz, R., and Ruiz-Carrillo, A. (1990c). Replacement Of Histone H1 By H5 In vivo Does Not Change the Nucleosome Repeat Length Of Chromatin But Increases Its Stability. EMBO J. *9*, 1651-1658.

Sun, J. M., Wiaderkiewicz, R., and Ruizcarrillo, A. (1989). Histone H5 In the Control Of DNA-Synthesis and Cell-Proliferation. Science *245*, 68-71.

Sung, M. T. (1977). Phosphorylation and Dephosphorylation of Histone V (H5): Controlled Condenstation of Avian Erythrocyte Chromatin. Biochemistry *16*, 286-291.

Sutcliffe, J. S., Nakao, M., Christian, S., ôrstavik, K. H., Tommerup, N., Ledbetter, D. H., and Beaudet, A. L. (1994). Deletions of a differentially methylated CpG island at the SNRPN gene define a putative imprinting control region. Nat. Genet. *8*, 52-58.

Svendsen, C. N. and Smith, A. G. (1999). New prospects for human stem-cell therapy in the nervous system. Trends Neurosci. *22*, 357-364.

Taga, T. (1997). The signal transducer gp130 is shared by interleukin-6 family of haematopoietic and neurotrophic cytokines. Ann. Med. *29*, 63-72.

Tagarro, I., Wiegant, J., Raap, A. K., Gonzalez-Aguilera, J. J., and Fernandez-Peralta, A. M. (1994). Assignment of human satellite 1 DNA as revealed by fluorescent in situ hybridization with oligonucleotides. Hum. Genet. *93*, 125-128.

Takami, Y. and Nakayama, T. (1997). A single copy of linker H1 genes is enough for proliferation of the DT40 chicken B cell line, and linker H1 variants participate in regulation of gene expression. Genes Cells *2*, 711-723.

Taylor, S. M. and Jones, P. A. (1979). Multiple New Phenotypes Induced in 10T1/2 and 3T3 Cells treated with 5-Azacytidine. Cell *17*, 771-779.

Tazi, J. and Bird, A. (1990). Alternative Chromatin Structure At CpG Islands. Cell *60*, 909-920.

Ten Hagen, K. G. and Cohen, S. N. (1993). Timing of replication of β-satellite repeats of human chromosomes. Nucleic Acids Res. *21*, 2139-2142.

Thoma, F. and Koller, T. (1981). Unravelled Nucleosomes, Nucleosome Beads and Higher Order Structures of Chromatin: Influence of Non-histone Components and Histone H1. J. Mol. Biol. *149*, 709-733.

Thoma, F., Koller, T. H., and Klug, A. (1979). Involvement of Histone H1 in the Organisation of the Nucleosome and of the Salt-Dependent Superstructures of Chromatin. J. Cell Biol. *83*, 403-427.

Thoma, F. and Koller, Th. (1977). Influence of Histone H1 on Chromatin Structure. Cell *12*, 101-107.

Thomas, G. and Padayatty, J. D. (1984). The Eukaryotic Gene. Curr. Sci. *53*, 1274-1279.

Thomas, J. O. (1999). Histone H1: location and role. Curr. Opin. Cell Biol. *11*, 312-317.

Thomas, J. O. and Rees, C. (1983). Exchange Of Histone-H1 and Histone-H5 Between Chromatin Fragments - a Preference Of H-5 For Higher-Order Structures. Eur. J. Biochem. *134*, 109-115.

Thomas, J. O. and Wilson, C. M. (1986). Selective radiolabelling and identification of a strong nucleosome binding site on the globular domain of histone H5. EMBO J. *5*, 3531-3537.

Thomas, K. R. and Capecchi, M. R. (1987). Site-Directed Mutagenesis by Gene Targeting in Mouse Embryo-Derived Stem Cells. Cell *51*, 503-512.

Thomson, J. A., Itskovitz-Eldor, J., Shapiro, S. S., Waknitz, M. A., Swiergiel, J. J., Marshall, V. S., and Jones, J. M. (1998). Embryonic stem cell lines derived from human blastocysts. Science *282*, 1145-1147.

Thomson, S., Clayton, A. L., Hazzalin, C. A., Rose, S., Barratt, M. J., and Mahadevan, L. C. (1999). The nucleosomal response associated with immediate-early gene induction is mediated via alternative MAP kinase cascades: MSK1 as a potential histone H3/HMG-14 kinase. EMBO J. *18*, 4779-4793.

Tonjes, R. R., Paul, D., and Doenecke, D. (1997). Transgenic Mice Transcribing the Human H1° Histone Gene Exhibit a Normal Phenotype. Eur. J. Biochem. *245*, 97-102.

Travers, A. (1999). The location of the linker histone on the nucleosome. Trends Biochem. Sci. *24*, 4-7.

Trieschmann, L., Martin, B., and Bustin, M. (1998). The chromatin unfolding domain of chromosomal protein HMG-14 targets the N-terminal tail of histone H3 in nucleosomes. Proc. Natl. Acad. Sci. U. S. A *95* , 5468-5473.

Tse, C., Sera, T., Wolffe, A. P., and Hansen, J. C. (1998). Disruption of higher-order folding by core histone acetylation dramatically enhances transcription of nucleosomal arrays by RNA polymerase III. Mol. Cell. Biol. *18*, 4629-4638.

Tumbar, T., Sudlow, G., and Belmont, A. S. (1999). Large-scale chromatin unfolding and remodeling induced by VP16 acidic activation domain. J. Cell Biol. *145*, 1341-1354.

Tung, W. L. and Chow, K. C. (1995). A modified medium for efficient electrotransformation of E. coli. Trends Genet. *11*, 128-129.

Turner, B. M., Birley, A. J., and Lavender, J. (1992). Histone H4 isoforms acetylated at specific lysine residues define individual chromosomes and chromatin domains in Drosophila polytene nuclei. Cell *69* , 375-384.

Turner, B. M. and Fellows, G. (1989). Specific antibodies reveal ordered and cell-cycle-related use of histone-H4 acetylation sites in mammalian cells. Eur. J. Biochem. *179*, 131-139.

Tyler-Smith, C., Gimelli, G., Giglio, S., Floridia, G., Pandya, A., Terzoli, G., Warburton, P. E., Earnshaw, W. C., and Zuffardi, O. (1999). Transmission of a fully functional human neocentromere through three generations. Am. J. Hum. Genet. *64*, 1440-1444.

Udvardy, A. (1999). Dividing the empire: boundary chromatin elements delimit the territory of enhancers. EMBO J. *18*, 1-8.

Umesono, K., Toda, T., Hayashi, S., and Yanagida, M. (1983). Cell division cycle genes nda2 and nda3 of the fission yeast Schizosaccharomyces pombe control microtubular organization and sensitivity to anti-mitotic benzimidazole compounds. J. Mol. Biol. *168*, 271-284.

Unni, E., Zhang, Y., Kangasniemi, M., Saperstein, W., Moss, S. B., and Meistrich, M. L. (1995). Stage-specific distribution of the spermatid-specific histone 2B in the rat testis. Biol. Reprod. *53*, 820-826.

Ura, K., Kurumizaka, H., Dimitrov, S., Almouzni, G., and Wolffe, A. P. (1997). Histone acetylation: Influence on transcription, nucleosome mobility and positioning, and linker histone-dependent transcriptional repression. EMBO J. *16*, 2096-2107.

Ura, K., Nightingale, K., and Wolffe, A. P. (1996). Differential association of HMG1 and linker histones B4 and H1 with dinucleosomal DNA: structural transitions and transcriptional repression. EMBO J. *15* , 4959-4969.

Ushinsky, S. C., Bussey, H., Ahmed, A. A., Wang, Y., Friesen, J., Williams, B. A., and Storms, R. K. (1997). Histone H1 in Saccharomyces cerevisiae. Yeast *13*, 151-161.

Utley, R. T., Ikeda, K., Grant, P. A., Cote, J., Steger, D. J., Eberharter, A., John, S., and Workman, J. L. (1998). Transcriptional activators direct histone acetyltransferase complexes to nucleosomes. Nature *394*, 498-502.

Vaccari, T., Beltrame, M., Ferrari, S., and Bianchi, M. E. (1998). HMG4, a new member of the HMG1/2 gene family. Genomics *49* , 247-252.

Van Holde, K. E. (1988). Chromatin. (New York: Springer Verlag).

Van Holde, K. E. and Zlatanova, J. (1996). What Determines the Folding Of the Chromatin Fiber. Proc. Natl. Acad. Sci. U. S. A. *93*, 10548-10555.

Van Sloun, P. P., Jansen, J. G., Weeda, G., Mullenders, L. H., van Zeeland, A. A., Lohman, P. H., and Vrieling, H. (1999). The role of nucleotide excision repair in protecting embryonic stem cells from genotoxic effects of UV-induced DNA damage. Nucleic Acids Res. *27*, 3276-3282.

van Steensel, B. and de Lange, T. (1997). Control of telomere length by the human telomeric protein TRF1. Nature *385*, 740-743.

van Steensel, B., Smogorzewska, A., and de Lange, T. (1998). TRF2 protects human telomeres from end-to-end fusions. Cell *92*, 401-413.

Vasioukhin, V., Degenstein, L., Wise, B., and Fuchs, E. (1999). The magical touch: genome targeting in epidermal stem cells induced by tamoxifen application to mouse skin. Proc. Natl. Acad. Sci. U. S. A. *96*, 8551-8556.

Verreault, A., Kaufman, P. D., Kobayashi, R., and Stillman, B. (1996). Nucleosome assembly by a complex of CAF-1 and acetylated histones H3/H4. Cell *87*, 95-104.

Vettese-Dadey, M., Grant, P. A., Hebbes, T. R., C, Allis, C. D., and Workman, J. L. (1996). Acetylation of histone H4 plays a primary role in enhancing transcription factor binding to nucleosomal DNA in vitro. EMBO J. *15*, 2508-2518.

Vincent, S., Marty, L., and Fort, P. (1993). S26 ribosomal protein RNA: an invariant control for gene regulation experiments in eukaryotic cells and tissues. Nucleic Acids Res. *21*, 1498.

Vissel, B. and Choo, K. H. (1987). Human Alpha-Satellite DNA - Consensus Sequence and Conserved Regions. Nucleic Acids Res. *15*, 6751-6752.

Vissel, B. and Choo, K. H. (1989). Mouse Major (Gamma) Satellite DNA Is Highly Conserved and Organized Into Extremely Long Tandem Arrays - Implications For Recombination Between Nonhomologous Chromosomes. Genomics *5*, 407-414.

Voullaire, L. E., Slater, H. R., Petrovic, V., and Choo, K. H. A. (1993). A Functional Marker Centromere with No Detectable α-Satellite, Satellite III, or CENP-B Protein: Activation of a Latent Centromere? Am. J. Hum. Genet. *52*, 1153-1163.

Wakayama, T., Perry, A. C., Zuccotti, M., Johnson, K. R., and Yanagimachi, R. (1998). Full-term development of mice from enucleated oocytes injected with cumulus cell nuclei. Nature *394*, 369-374.

Wandall, A., Tranebjaerg, L., and Tommerup, N. (1998). A neocentromere on human chromosome 3 without detectable α-satellite DNA forms morphologically normal kinetochores. Chromosoma *107*, 359-365.

Wang, Z. F., Krasikov, T., Frey, M. R., Wang, J., Matera, A. G., and Marzluff, W. F. (1996a). Characterization of the mouse histone gene cluster on chromosome 13: 45 histone genes in three patches spread over 1Mb. Genome Res. *6*, 688-701.

Wang, Z. F., Sirotkin, A. M., Buchold, G. M., Skoultchi, A. I., and Marzluff, W. F. (1997). The mouse histone H1 genes: Gene organization and differential regulation. J. Mol. Biol. *271*, 124-138.

Wang, Z. F., Tisovec, R., Debry, R. W., Frey, M. R., Matera, A. G., and Marzluff, W. F. (1996b). Characterization of the 55-kb mouse histone gene cluster on chromosome 3. Genome Res. *6*, 702-714.

Warburton, P. E., Cooke, C. A., Bourassa, S., Vafa, O., Sullivan, B. A., Stetten, G., Gimelli, G., Warburton, D., Tylersmith, C., Sullivan, K. F., Poirier, G. G., and Earnshaw, W. C. (1997). Immunolocalization of CENP-A suggests a distinct nucleosome structure at the inner kinetochore plate of active centromeres. Curr. Biol. *7*, 901-904.

Warburton, P. E. and Cooke, H. J. (1997). Hamster chromosomes containing amplified human α-satellite DNA show delayed sister chromatid separation in the absence of de novo kinetochore formation. Chromosoma *106*, 149-159.

Waye, J. S., Mitchell, A. R., and Willard, H. F. (1988). Organization and genomic distribution of "82H" α-satellite DNA. Hum. Genet.

Waye, J. S. and Willard, H. F. (1989). Human β -satellite DNA: genomic organization and sequence definition of a class of highly repetitive tandem DNA. Proc. Natl. Acad. Sci. U. S. A. *86*, 6250-4.

Wei, Y., Yu, L. L., Bowen, J., Gorovsky, M. A., and Allis, C. D. (1999). Phosphorylation of histone H3 is required for proper chromosome condensation and segregation. Cell *97*, 99-109.

Weier, H. U. G., Kleine, H. D., and Gray, J. W. (1991). Labeling of the Centromeric Region On Human Chromosome-8 By In situ Hybridization. Hum. Genet. *87*, 489-494.

Weintraub, H. (1978). The nucleosome repeat length increases during erythropoiesis in the chick. Nucleic Acids Res. *5*, 1179-1188.

Weintraub, H. and Groudine, M. (1976). Chromosomal subunits in active genes have an altered conformation. Science *193*, 848-856.

Weisbrod, S. and Weintraub, H. (1979). Isolation of a subclass of nuclear proteins responsible for conferring a DNase I-sensitive structure on globin chromatin. Proc. Natl. Acad. Sci. U. S. A *76*, 630-634.

Weiss, M. J. (1997). Embryonic stem cells and hematopoietic stem cell biology. Hematol. Oncol. Clin. North Am. *11*, 1185-1198.

Weiss, M. J. and Orkin, S. H. (1996). In vitro differentiation of murine embryonic stem cells. New approaches to old problems. J. Clin. Invest. *97*, 591-595.

Whitehouse, I., Flaus, A., Cairns, B. R., White, M. F., Workman, J. L., and Owen-Hughes, T. (1999). Nucleosome mobilization catalysed by the yeast SWI/SNF complex. Nature *400*, 784-787.

Widom, J. (1997). Chromatin: the nucleosome unwrapped. Curr. Biol. *7*, R653-R655.

Widom, J., Finch, J. T., and Thomas, J. O. (1985). Higher-Order Structure of Long Repeat Chromatin. EMBO J. *4*, 3189-3194.

Wiekowski, M., Miranda, M., Nothias, J. Y., and DePamphilis, M. L. (1997). Changes in histone synthesis and modification at the beginning of mouse development correlate with the establishment of chromatin mediated repression of transcription. J. Cell Sci. *110*, 1147-1158.

Willard, H. F. (1998). Centromeres: the missing link in the development of human artificial chromosomes. Curr. Opin. Genet. Dev. *8*, 219-225.

Williams, B. C., Murphy, T. D., Goldberg, M. L., and Karpen, G. H. (1998). Neocentromere activity of structurally acentric mini-chromosomes in Drosophila. Nat. Genet. *18*, 30-37.

Williams, R. L., Hilton, D. J., Pease, S., Willson, T. A., Stewart, C. L., Gearing, D. P., Wagner, E. F., Metcalf, D., Nicola, N. A., and Gough, N. M. (1988). Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. Nature *336*, 684-687.

Williams, S. P., Athey, B. D., Muglia, L. J., Schappe, R. S., Gough, A. H., and Langmore, J. P. (1986). Chromatin fibers are left-handed double helices with diameter and mass per unit length that depend on linker length. Biophys. J. *49*, 233-248.

Wilmut, I., Schnieke, A. E., McWhir, J., Kind, A. J., and Campbell, K. H. (1997). Viable offspring derived from fetal and adult mammalian cells. Nature *385*, 810-813.

Wittschieben, B. O., Otero, G., de Bizemont, T., Fellows, J., Erdjument-Bromage, H., Ohba, R., Li, Y., Allis, C. D., Tempst, P., and Svejstrup, J. Q. (1999). A novel histone acetyltransferase is an integral subunit of elongating RNA polymerase II holoenzyme. Mol. Cell *4*, 123-128.

Wobus, A. M. and Guan, K. M. (1998). Embryonic stem cell-derived cardiac differentiation: Modulation of differentiation and "loss-of-function" analysis in vitro. Trends Cardiovasc. Med. *8*, 64-74.

Wolffe, A. P. (1995). Chromatin Structure and Function. (London: Academic Press).

Wolffe, A. P. (1996). Histone deacetylase: a regulator of transcription. Science *272*, 371-372.

Wolffe, A. P. (1999). Architectural regulations and HMG1. Nat. Genet. *22*, 215-217.

Wolffe, A. P. and Hayes, J. J. (1999). Chromatin disruption and modification. Nucleic Acids Res. *27*, 711-720.

Wong, A. K. C., Biddle, F. G., and Rattner, J. B. (1990). The Chromosomal Distribution of the Major and Minor Satellite Is Not Conserved in the Genus Mus. Chromosoma *99*, 190-195.

Wong, A. K. C. and Rattner, J. B. (1988). Sequence Organization and Cytological Localization of the Minor Satellite of Mouse. Nucleic Acids Res. *16*, 11645-11661.

Woodcock, C. L. (1994). Chromatin fibers observed in situ in frozen hydrated sections. Native fiber diameter is not correlated with nucleosome repeat length. J. Cell Biol. *125*, 11-19.

Woodcock, C. L., Frado, L. L. Y., and Rattner, J. B. (1984). The Higher-Order Structure of Chromatin - Evidence For a Helical Ribbon Arrangement. J. Cell Biol. *99*, 42-52.

Woodcock, C. L., Grigoryev, S. A., Horowitz, R. A., and Whitaker, N. (1993). A chromatin folding model that incorporates linker variability generates fibers resembling the native structures. Proc. Natl. Acad. Sci. U. S. A. *90*, 9021-9025.

Woodcock, C. L. and Horowitz, R. A. (1995). Chromatin organization re-viewed. Trends Cell Biol. *5*, 272-277.

Workman, J. L. and Kingston, R. E. (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. Ann. Rev. Biochem. *67*, 545-579.

Wreggett, K. A., Hill, F., James, P. S., Hutchings, A., Butcher, G. W., and Singh, P. B. (1994). A mammalian homologue of Drosophila heterochromatin protein 1 (HP1) is a component of constitutive heterochromatin. Cytogenet. Cell Genet. *66*, 99-103.

Wu, R. S. and Bonner, W. M. (1981). Separation of basal histone synthesis from S-phase histone synthesis in dividing cells. Cell *27*, 321-330.

Wu, R. S., Panusz, H. T., Hatch, C. L., and Bonner, W. M. (1986). Histones and Their Modifications. CRC Crit. Rev. Biochem. *20*, 201-263.

Xie, X., Kokubo, T., Cohen, S. L., Mirza, U. A., Hoffman, A., Chait, B. T., Roeder, R. G., Nakatani, Y., and Burley, S. K. (1996). Structural similarity between TAFs and heterochimeric core of the histone octamer. Nature *380*, 316-322.

Yang, G., Leuba, S. H., Bustamante, C., Zlatanova, J., and van Holde, K. (1994). Role of linker histones in extended chromatin fibre structure. Nat. Struct. Biol. *1*, 761-763.

Yao, J., Lowary, P. T., and Widom, J. (1990a). Direct detection of linker DNA bending in defined-length oligomers of chromatin. Proc. Natl. Acad. Sci. U. S. A. *87*, 7603-7607.

Yao, J., Lowary, P. T., and Widom, J. (1991). Linker DNA bending induced by the core histones of chromatin. Biochemistry *30*, 8408-8414.

Yoda, K., Ando, S., Okuda, A., Kikuchi, A., and Okazaki, T. (1998). In vitro assembly of the CENP-B/α-satellite DNA/core histone complex: CENP-B causes nucleosome positioning. Genes Cells *3*, 533-548.

Yoshida, M., Hornouchi, S., and Beppu, T. (1995). Trichostatin A and trapoxin: novel chemical probes for the role of histone acetylation in chromatin structure and function. Bioessays *17*, 423-430.

Zeiger, R. S., Salomon, R., and Peacock, A. C. (1971). Isolation of mouse satellite deoxyribonucleic acid by composite polyacrylamide gel electrophoresis. Biochemistry *10*, 4219-4223.

Zentgraf, H. and Franke, W. W. (1984). Differences of supranucleosomal organization in different kinds of chromatin: cell type-specific globular subunits containing different numbers of nucleosomes. J. Cell Biol. *99*, 272-286.

Zhang, X. Y., Fittler, F., and Horz, W. (1983). 8 Different Highly Specific Nucleosome Phases On α-Satellite DNA in the African-Green Monkey. Nucleic Acids Res. *11*, 4287-4306.

Zhang, X. Y. and Horz, W. (1982). Analysis of Highly Purified Satellite DNA Containing Chromatin From the Mouse. Nucleic Acids Res. *10*, 1481-1494.

Zhang, X. Y. and Horz, W. (1984). Nucleosomes are Positioned on Mouse Satellite DNA in Multiple Highly Specific Frames that are Correlated with a Diverged Subrepeat of Nine Base-pairs. J. Mol. Biol. *176*, 105-129.

Zhang, Y., Ng, H. H., Erdjument-Bromage, H., Tempst, P., Bird, A., and Reinberg, D. (1999). Analysis of the NuRD subunits reveals a histone deacetylase core complex and a connection with DNA methylation. Genes Dev. *13*, 1924-1935.

Zhao, K., Hart, C. M., and Laemmli, U. K. (1995). Visualization of chromosomal domains with boundary element-associated factor BEAF-32. Cell *81*, 879-889.

Zhao, K., Kas, E., Gonzalez, E., and Laemmli, U. K. (1993). SAR-dependent mobilization of histone H1 by HMG-I/Y in vitro: HMG-I/Y is enriched in H1-depleted chromatin. EMBO J. *12*, 3237-3247.

Zhao, K. J., Wang, W. D., Rando, O. J., Xue, Y. T., Swiderek, K., Kuo, A., and Crabtree, G. R. (1998). Rapid and phosphoinositol-dependent binding of the SWI/SNF-like BAF complex to chromatin after T lymphocyte receptor signaling. Cell *95*, 625-636.

Zhou, Y. B., Gerchman, S. E., Ramakrishnan, V., Travers, A., and Muyldermans, S. (1998). Position and orientation of the globular domain of linker histone H5 on the nucleosome. Nature *395*, 402-405.

Zink, D., Cremer, T., Saffrich, R., Fischer, R., Trendelenburg, M. F., Ansorge, W., and Stelzer, E. H. (1998). Structure and dynamics of human interphase chromosome territories in vivo. Hum. Genet. *102*, 241-251.

Zlatanova, J. and Van Holde, K. E. (1998). Linker histones versus HMG1/2: a struggle for dominance? Bioessays *20*, 584-588.

Zuckerkandl, E. (1997). Junk DNA and sectorial gene repression. Gene *205*, 323-343.